

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 23, 2003, 17:44:27 : Search time 3326 Seconds
(without alignments)
3202.534 Million cell updates/sec

Title: US-09-842-364-3
Perfect score: 366
Sequence: 1 MASMAVLTWALALSAFSA.....LMEDITHSLDGHSHLGD 366

Scoring table:
ORIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 6 aa = 18 nt

Total number of hits satisfying chosen parameters: 175711

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODL=frame_plus_p2n_model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09842364/runat.16012003.164829.10649/app.query.fasta.1.519
-DB=genembl -OEMT=fastap -SUFFIX=olig.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT-0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=6 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09842364 -RCGN_1_1_2425 -runat_16012003.164829.10649 -NCPU=6 -ICPU=3
-NO_XLAPY -NO_MMAL -LARGOQUERY -NFC_SCORES=0 -WAIT -IONGLOC -DEV_TIMPOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-Ygapop=60 -Ygapext=60 -DELOP=6 -DELEXT=7

Database: Genembl:
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

Result No. Score Match Length DB ID

Result No.	Score	Match Length	DB	ID	Description
1	366	100.0	1322	6	AX06773 Sequence
2	366	100.0	1349	9	AF202890 Homo sapi
3	366	100.0	1867	6	AX006774 Sequence
4	366	100.0	1879	6	AX063165 Sequence
5	366	100.0	1879	6	AX469877 Sequence
6	366	100.0	1889	6	AF202889
7	366	100.0	1894	6	AX061634
8	312	85.2	5381	6	AX063467 Sequence
9	312	85.2	5381	6	AX469879 Sequence
10	312	85.2	42479	6	AC074203
11	312	85.2	81001	6	AX063464
12	312	85.2	81001	6	AX469876 Sequence
13	312	85.2	115027	2	AF001480
14	312	85.2	115582	2	AF001481
15	312	85.2	188971	9	AC007707
16	279	76.2	1371	6	AX201372
17	279	76.2	1371	6	AX073667
18	213	58.2	1334	6	AC113242
19	196	53.6	181412	2	AC113242
20	176	48.1	592	2	AF129257
21	166	44.0	191656	2	AC118574
22	166	44.0	191656	2	BC011198
23	166	44.0	191656	2	AF327059
24	166	44.0	191656	2	AF401201
25	166	44.0	191656	2	AC118580
26	166	44.0	191656	2	AX006770
27	166	44.0	191656	2	AF202888
28	166	44.0	191656	2	AX006771
29	166	44.0	191656	2	AF202887
30	166	44.0	191656	2	AC118580
31	166	44.0	191656	2	AX323765 Sequence
32	166	44.0	191656	2	AC111620
33	166	44.0	191656	2	YSC70B3
34	166	44.0	191656	2	YSC70B3
35	166	44.0	191656	2	YSC70B3
36	166	44.0	191656	2	YSC70B3
37	166	44.0	191656	2	YSC70B3
38	166	44.0	191656	2	YSC70B3
39	166	44.0	191656	2	YSC70B3
40	166	44.0	191656	2	YSC70B3
41	166	44.0	191656	2	YSC70B3
42	166	44.0	191656	2	YSC70B3
43	166	44.0	191656	2	YSC70B3
44	166	44.0	191656	2	YSC70B3
45	166	44.0	191656	2	YSC70B3

RESULT 1

ALIGNMENTS

AX006773
 LOCUS AX006773 1322 bp DNA linear PAT 06-SEP-2000
 DEFINITION Sequence 22 from Patent WO003013.
 ACCESSION AX006773
 VERSION AX006773.1 GI:9994814
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1322)
 Leegwater, A.C., Van der Vijet, H.N., Chamuleau, R.A. and Groenink, M.
 Gene and protein involved in liver regeneration
 JOURNAL. patient: WO 0003013-A 22 20-JAN-2000.
 LEGGATER ADAM CORNELIS JOZEF (NL); VIJET HENDRIK NIELS V D (NL);
 AMSTERDAM MOLECULAR THERAPEUTI (NL); CHAMULEAU ROBERT ANTOINE FRANC
 (NL); GROENINK MARTIN (NL)
 FEATURES
 source 1..1322
 location/Qualifiers
 BASE COUNT 267 a 408 c 427 g 220 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 length: 1322
 Score: 366.00 Matches: 366
 Percent Similarity: 100.00% Conservation: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-842-364-3 (1-366) x AX006773 (1-1322)
 QY 1 MetlaserMetlalaValleuthrTrpAlaLeuSerAla 20
 Db 10 ATGCCAACATCGCTGCGCTCCTCACCAGCGCTTCCTTCACGCTTTCCGCC 69
 QY 21 ThrGlnAlaArgLysGlyPheTrpAspTyrPheSerGlnThrSerGlyAspLysArg 40
 Db 70 ACCCAGGACGCGAAAGCTTCTGGAGCTACTTACGCCAGCAGCGGGGACAAAGGCG 129
 QY 41 ValGlnGlnIleHisGlnGlnIleLysMetAlaArgGluProAlaThrIleLysAspSerLeu 60
 Db 130 GTGGAGCGATTCATCAGCAGAAAGATGCTCGCGAGCCCGGACCTTGAAGACAGCCTT 189
 QY 61 GlnGlnAspLeuAsnAsnMetLanLysPheLeuGluLysLeuArgProLeuSerGlySer 80
 Db 190 GAGCAAGACCTTCAACAAATATGAACAGATTCTCGCAAAAGCTCAGCCCTTGAGTGGAGC 249
 QY 81 GlnAlaProAlaArgLeuProGlnAspProValGlyMetLarArgGlnLeuGlnLeuLeu 100
 Db 250 GAGGCTCTCGGCTCCACAGAGACCGGCTGGGCGATCGCGGCGGACCTGACAGAGAGTTG 309
 QY 101 GlnGlnValLysAlaArgLeuGlnProTyrMetAlaGlnAlaHisGlnLeuValGlyTyr 120
 Db 310 GAGGAGGTGAAGGCTCGCTCCAGCCCTTACATGGCAGAGCGCAGAGAGCTGGTGGCTGG 369
 QY 121 AsnLeuGlnGlyLeuArgGlnGlnLeuLysProTyrThrMetLanLeuMetGlnGlnVal 140
 Db 370 AATTTGGAGGCTTCGCGCAGCACTGAAAGCCCTTACACCATGATGTATGACAGAGGTG 429
 QY 141 AlaLeuArgValGlnGlnLeuGlnGlnLeuArgValValGlyGlnAspThrLysAla 160
 Db 430 GCCCTGGCGCGTGCAGAGAGCTGAGAGCAGTTCGCGGTGGGGGAGAGACCAAGGCC 489
 QY 161 GlnLeuLeuGlnGlyGlyValAspGlnAlaTrpAlaLeuLeuGlnGlnGlnSerArgVal 180
 Db 490 CAGTTGGCTGGGGGGGTGAGAGAGCTTGGCTGGTGCAGGAGCTCAGAGCGCGCTG 549
 QY 181 ValHisHisThrGlyArgPheLysGlnLeuPheHisProTyrAlaGlnSerLeuValSer 200
 Db 550 GTGCACACACCGCGCGCTTCAAGAGCTTTCACCCCATTCAGCCGAGAGCTGGTGAGC 609

QY 201 GlnIleGlyArgHisValGlnGlnLeuHisArgSerValAlaProHisAlaProAlaSer 220
 Db 610 GGCATCGGGCCGACGTGACAGAGCTGCACCGCATGTGGCTCCGACAGCCGCCAGC 669
 QY 221 ProAlaArgLeuSerArgCysValGlnValLeuSerArgLysIleuThrLeuLysAlaLys 240
 Db 670 CCCCGCCGCTCAGTCGCTGCGCTGCGAGCTGCTCCCGAAGCTACCTCAGCGCCAAAG 729
 QY 241 AlaLeuHisAlaArgIleGlnGlnAsnLeuAspGlnLeuArgGlnGlnLeuSerArgAla 260
 Db 730 GCCCTGCACGACCATGACATGACAGCAAACTGCAGCAGCTCCGCGAAGACTCAGCAGACC 789
 QY 261 PheAlaGlyThrGlnGlyThrGlnGlnGlnValGlnGlyProAspProGlnMetLeuSerGlnGlu 280
 Db 790 TTGGAGCAGCTGGGAGCTGAGAGAGGGCCGCGCGGACCCGACATGCTCTCCGAGAG 849
 QY 281 ValArgGlnArgLeuGlnAlaPheArgGlnAspThrTrpLeuGlnIleAlaAlaPheThr 300
 Db 850 GTCCGCGACGACCTTCAGAGCTTTCGCCGACGACACTTACCTGACATGAGCTTCACT 909
 QY 301 ArgAlaIleAspGlnGlnThrGlnGlnValGlnGlnGlnLeuAlaProProProGly 320
 Db 910 CCGCCATGACACAGACACTGACAGAGCTCCACGACAGCTCCGCGACCTCCACAGC 969
 QY 321 HisSerAlaPheAlaProGlnPheGlnGlnThrAspSerGlyLysValLeuSerLysLeu 340
 Db 970 CACAGTGCCTTGGCCCGCAGAGTTTCACACAAACAGACAGTGGCAGAGCTTGACGAACTG 1029
 QY 341 GlnAlaArgLeuAspAspLeuTrpClnuAspIleThrHisSerLeuHisAspGlnGlnHis 360
 Db 1030 CAGCCCGCTGATGACCTGCGGAGAGATCAGTACAGCAGCTTCATACAGAGCGCCAC 1089
 QY 361 SerHisLeuGlyAspPro 366
 Db 1090 ACCCATTCGGGGACCC 1107
 RESULT 2
 LOCUS AF202890 1349 bp mRNA linear PRI 12-APR-2002
 DEFINITION Homo sapiens apolipoprotein A5 (APOA5) mRNA, complete cds,
 alternative transcript.
 ACCESSION AF202890
 VERSION AF202890.1 GI:6707434
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1349)
 Pennacchio, L.A., Olivier, M., Hubacek, J.A., Cohen, J.C., Cox, D.R.,
 Fruchart, J.C., Krauss, R.M. and Rubin, E.M.
 An apolipoprotein influencing triglycerides in humans and mice
 revealed by comparative sequencing
 JOURNAL. Science 294 (5540), 169-173 (2001)
 MEDLINE 21472646
 PUBMED 11588264
 REFERENCE
 2 (bases 1 to 1349)
 van der Vijet, H.N., Samuels, M.G., Leegwater, A.C., Levels, J.H.,
 Reitsma, P.H., Boers, W. and Chamuleau, R.A.
 Apolipoprotein A-V: a novel apolipoprotein associated with an early
 phase of liver regeneration
 JOURNAL. J. Biol. Chem. 276 (48), 44512-44520 (2001)
 MEDLINE 21576234
 PUBMED 11577099
 REFERENCE
 3 (bases 1 to 1349)
 Van der Vijet, H.N., Groenink, M., Leegwater, A.C.J. and
 Chamuleau, R.A.F.M.
 Direct Submission
 TITLE Submitted (09-NOV-1999) Experimental Hepatology, Academic Medical
 Center, Meibergdreef 9, Amsterdam 1105 AZ, Netherlands
 FEATURES
 source 1..1349

QY 201 GYITIEGLYArgHisValGlnGlnLeuHisAsgSerValAlaProHisAlaProAlaSer 220
 Db 621 GGCATCGGGCCCGCCAGCGAGAGCTGCACCCAGCTGCTCCGACCCCGCCGAC 680
 QY 221 ProAlaAsgLeuSerAsgCysValGlnValLeuSerAsgLysLeuThreLeuLysAlaLys 240
 Db 681 CCGGGGCGGCTCAGTCGCTGCTGCGAGGTGCTCTCCGGAGCTCAGCTCAGGCCAAG 740
 QY 241 AlaLeuHisAlaAArgIleGlnGlnAsnLeuAsgLysLeuAsgLysLeuSerAsgAla 260
 Db 741 GCGCTGACAGCAGCATCCAGAGAACCTGCAGCAGCTGCGGAGAGCTCAGCAGAGCC 800
 QY 261 PheAlaGlyThrGlyThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 Db 801 TTTCCAGGACACTGAGCTCAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860
 QY 281 ValAsgGlnAsgLeuGlnAlaAAspThrGlnAsgThrGlnGlnGlnGlnGlnGlnGln 300
 Db 861 GTGGCCAGCGACCTCAGGCTTCCGCCAGGACACCTACCTCAGATGAGCTGCTTCACT 920
 QY 301 ArgAlaIleAsgPngInGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
 Db 921 CGCGCCATCGACAGAGAGCTGAGAGGTCCAGCAGCAGCTGCGCCACCTCCACAGGC 980
 QY 321 HisSerAlaPheAlaProGlnPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 340
 Db 981 CACACTGCTTCCGCGCCAGAGTTTCACACACACAGCAGTGGCAGGTTCTGACCAAGCTG 1040
 QY 341 GlnAlaAsgLeuAsgPngLysPngLysPngLysPngLysPngLysPngLysPngLys 360
 Db 1041 CAGGCCCTCTGATGAGCTGTGGAGAGCATCACTACAGCTTCTCATGACAGGCGCCAC 1100
 QY 361 SerHisLeuGlyAsgPro 366
 Db 1101 AGCCATCTGGGGGACCC 1118

RESULT 6
 AF202889
 LOCUS Homo sapiens apolipoprotein A5 (APOA5) mRNA, complete cds,
 DEFINITION alternative transcript.
 AF202889
 ACCESSION AF202889.1 GI:6707432
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 1889)
 Pennachio, L.A., Olivier, M., Hubacek, J.A., Cohen, J.C., Cox, D.R.,
 Fruchart, J.C., Krauss, R.M., and Rubin, E.M.
 An apolipoprotein influencing triglycerides in humans and mice
 revealed by comparative sequencing
 Science 294 (5540), 169-173 (2001)
 2 (bases 1 to 1889)
 van der Vliet, H.N., Samuels, M.C., Leegwater, A.C., Levels, J.H.,
 Reitsma, P.H., Boers, W., and Chamulieu, R.A.
 Apolipoprotein A-V: a novel apolipoprotein associated with an early
 phase of liver regeneration
 J. Biol. Chem. 276 (48), 44512-44520 (2001)
 3 (bases 1 to 1889)
 van der Vliet, H.N., Groenink, M., Leegwater, A.C.J., and
 Chamulieu, R.A.F.M.
 Direct Submission
 Submitted (09-NOV-1999) Experimental Hepatology, Academic Medical
 Center, Meibergdreef 9, Amsterdam 1105 AZ, Netherlands
 FEATURES
 Source 1..1889

Alignment Scores:
 Pred. No.: 0 Length: 1889
 Score: 366.00 Matches: 366
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-842-364-3 (1-366) x AF202889 (1-1889)

QY 1 MetAlaSerMetAlaValLeuThrTrpAlaLeuAlaLeuSerAlaAAspSerAla 20
 Db 9 ATGGCAAGCATGGCTGCGCTGCTCACTGCTGCTGCTGCTTCTTTCACGCTTTCGCGC 68

QY 21 ThrGlnAlaAsgLysGlyPheThrPaspTyrPheSerGlnThrSerGlyAAspLysGlyArg 40
 Db 69 ACCCAGGCGGAAAGGCTTCTGGGAGTACTTCCAGCCAGCAGCGGGGACAAAGGCGAG 128

QY 41 ValGlnGlnIleHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
 Db 129 CTGACAGCATCCATCCAGCAGAGAGATGCTCCGACCCCGCCAGCCCTGAACACAGCTT 188

QY 61 GlnGlnAsgLeuAsgAAspMetAlaAspLysPheLeuGlnGlnGlnGlnGlnGlnGln 80
 Db 189 GAGCAAGACCTTACCATATGTGACCAAGTTCCTGGAAAGCTGAGGCTCTGAGTGGAGC 248

QY 81 GlnAlaProAsgLeuProGlnAsgProValGlyMetAsgArgGlnGlnGlnGlnGln 100
 Db 249 GAGGCTCTCGGCTCCACAGAGACCGGATGGGCGATGGGCGGCGACCTCAGAGAGAGTTG 308

QY 101 GlnGlnValLysAlaAsgLeuGlnProTyrMetAlaGlnAlaHisGlnLeuValGlyTyr 120
 Db 309 GAGCAGGTGAAGCTCTGCTCCAGCCCTACATCCAGCAGCGCCAGCAGCTGCTGGCTGG 368

QY 121 AsnLeuGlnGlyLeuAsgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
 Db 369 AATTGGAGGCTTCCGCGACCACTGAAGCCCTTACAGATGAGATGATGAGAGAGGTG 428

QY 141 AlaLeuAsgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
 Db 429 GCGCTGCGCGCAGAGAGCTCAGAGAGAGTGGCGCGTGGGGGAGAGACCAAGGCC 488

QY 161 GlnLeuLeuGlyGlyValAAspGlnAlaAAspLysAlaLeuGlnGlnGlnGlnGln 180
 Db 489 CAGTTGCTGGGGCGGTGAGAGAGCTTGGCTTTCGACGAGGACTTCACAGCCGCGGTG 548

QY 181 ValHisHisThrGlyArgPheLysGlnLeuPheHisProTyrAlaGlnSerLeuValSer 200

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue="liver"
 /note="Caucasian"
 1..1889
 /gene="APOA5"
 18..1109
 /gene="APOA5"
 /note="constituent of HDL; concentration elevated after
 partial hepatectomy; alternative polyadenylation; long
 transcript"
 /codon_start=1
 /product="apolipoprotein A5"
 /protein_id="AA25661.1"
 /db_xref="GI:6707433"
 /translation="MAAVITWVALALISAFSAQAKGFWDFYSQSGDKGRVEQIHQO
 KMAEPATLKDSLEKDIANNMKFLKLPISGSAPRLDQPVGMRROLEFEVKA
 RLQPYMAAEHLVGNLEGLRQQLKPYTMDEVALVQLEDELRYVGETRQQL
 GYVDKAMLLGLOSRYVHHTGRFELHPYAESLSVSGIRHVOELHRSVAFHASP
 ARLSRCVOVLSRKLTLKAKALHARLQNLIDLRELSRAFGTEEGADPPMLSE
 FVKORLQAFRODTYLOIAAPTRAIIDQTEFEVQOQLAPPQHSAPAFQOTDSKVL
 SKLQARLDDLMEDLTHSLHDOCHSHLGP"


```

OY      355  LeuHISASPGInGlyHISerHisLeuGlyAspPro 366
          |||||||
DB      3144  CTTCATGACCAAGGCCACACCATCTGGGGGACCCC 3179

RESULT 9
LOCUS      AX469879          5381 bp    DNA      linear    PAT 09-AUG-2002
DEFINITION Sequence 4 from Patent WO02053734.
ACCESSION  AX469879
VERSION     AX469879.1  GI:22205152
KEYWORDS
SOURCE      human.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1
AUTHORS    Bouguerel, I., Duclet, A., Clusel, C., Dumas, M.E., Yen-Potin, F.,
            Denison, B., Bihain, B., Bour, B., Ebbels-Reed, D. and Salter-Cid, L.
            Methods and compositions for inhibiting neoplastic cell growth
            Patent: WO 02053734-A 4 11-JUL-2002;
            GENSET SA (FR)

FEATURES
             Location/Qualifiers
             source          1..3381
                             /organism="Homo sapiens"
                             /db_xref="taxon:9606"
             misc_feature    1..918
                             /note="5' regulatory region"
             primer_bind     300..318
                             /note="17-42-319.mis"
             misc_binding    307..331
                             /bound_molety="17-42-319.probe"
             variation       319
                             /note="17-42-319 : polymorphic base C or T"
             primer_bind     320..338
                             /note="17-42-319.mis complement"
             exon            919..930
                             /number=1
             variation       1241
                             /note="17-39-343 : T in ref genbank AC007707"
             exon            1442..1498
                             /number=2
             variation       1447
                             /note="17-40-202 : G in ref genbank AC007707"
             exon            1613..1724
                             /number=3
             exon            2243..3940
                             /number=4
             primer_bind     3194..3212
                             /note="17-41-250.mis"
             misc_binding    3201..3225
                             /bound_molety="17-41-250.probe"
             variation       3213
                             /note="17-41-250 : polymorphic base C or T"
             primer_bind     3214..3232
                             /note="17-41-250.mis complement"
             misc_feature    3941..3381
                             /note="3' regulatory region"
BASE COUNT  1104 a 1434 c 1709 g 1132 t 2 others
ORIGIN
Alignment Scores:
Pred. No.:      0          Length:      5381
Score:          312.00     Matches:      312
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    85.25%      Indels:      0
DB:             6          Gaps:         0

US-09-842-364-3 (1-366) x AX469879 (1-5381)
OY      55  ThrLeuTAspSerLeuGluGlnAspLeuAsnMetAlaGlnLysPheLeuGluLysLeu 74
          |||||||
DB      2244  ACCCTGAAGACGACCTTGAGCAAGACCTCAACATATGACCAAGATTCTGGAAAGACTG 2303

```

```

OY      75  ArgProLeuSerGlySerGluAlaProArgLeuProGlnAspProValGlyMetArg 94
          |||||||
DB      2304  AGCCCTGAGTGGAGGACGAGCTCTCGCTCCACAGACCCGCTGGCATCCGCCG 2363

OY      95  GlnLeuGlnGluGluLeuGluGluValLysAlaArgLeuGlnProTyrMetAlaGlu 114
          |||||||
DB      2364  CAGCTGCAGAGAGAGTGGAGAGAGTGAAGGCTCGCTCCAGCTTACCTGAGAGAGGCG 2423

OY      115  HisGluLeuValGlyTTPAsnLeuGluGluLeuArgGlnGlnLeuLysProTyrMet 131
          |||||||
DB      2424  CACGAGCTGGTGGCTGGAAATTTGGAGGCTGGCGGACAGCACTGAAGCCCTACAGCATG 2483

OY      135  AspleuMetGluGlnValAlaLeuArgValGlnGluLeuGlnGlnLeuArgVal 154
          |||||||
DB      2484  GATCTGATGACAGCAGCTGGCCCTCGCCGACGAGCACTTCACAGCAGCTTCCCTGCTG 2543

OY      155  GlyGluAspThrLysAlaGlnLeuLeuGlyGlyValAlaGlnAlaTProAlaLeuLeuGln 174
          |||||||
DB      2544  GGGAGAGACACCAAGGCCAGTGTGCGGGGCGTGGAGAGAGGCTTGGCTTGGCTGCGAG 2603

OY      175  GlyLeuGlnSerArgValAlaHisThrGlyArgPheLysGluLeuPheHisProTyr 194
          |||||||
DB      2604  GGACTGCAGAGCGCGGTGGTGACACACCGCGCTTCAMAGACTCTTCACCCATAC 2663

OY      195  AlaGluSerLeuValSerGlyTLeGlyArgHisValGlnGlnLeuHisArgSerValAla 214
          |||||||
DB      2664  CCGGAGACCTCTGTCAGCGCATCGCGCCGCGCAGCTTCAGAGCTTCCAGCTTGGCTGCT 2723

OY      215  ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
          |||||||
DB      2724  CCGCAGCGCCCGCCGACGCCCGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2783

OY      235  LeuThrLeuLysAlaLysAlaLeuHisAlaArgLieglnGlnAsnLeuAspGlnLeuArg 254
          |||||||
DB      2784  CTCACGCTCAAGGCCAAGGCCCTGCACGACGACATCCAGCAAGAACTGGACAGCTGGCG 2843

OY      255  GluGluLeuSerArgAlaPheAlaGlyThrGlyThrGluGluGlyAlaGlyProAspPro 274
          |||||||
DB      2844  CAGAGCTCAGAGAGCCCTTTCACAGCACTGCGAGCTGCAAGAGCGCGCGCGCGAGCC 2903

OY      275  GlnMetLeuSerGluGluValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeu 294
          |||||||
DB      2904  CAGATGCTCTCCAGAGAGGTGGCGCAGGAGCTTACAGGCTTTCGCGAGAGACCTACTG 2963

OY      295  GlnIleAlaAlaPheThrArgAlaAlaAspGlnGluThrGluGluValGlnGlnGlnLeu 314
          |||||||
DB      2964  CAGATAGCTGCTCTCAGCTGCGCGCATTCGACAGGAGACTGAGAGAGTCCACAGCAGCTG 3023

OY      315  AlaProProProProGlyHisSerAlaPheAlaProGluPheGlnThrAspSerGly 334
          |||||||
DB      3024  GCGCCACCTCCACAGGCCACAGTCCCTTCCGCCCAAGTTTCAACAAACAGACAGTGGC 3083

OY      335  LysValLeuSerLysLeuGlnAlaArgLeuAspLeuTrpGluAspIleThrHisSer 354
          |||||||
DB      3084  AAGSTTCTGAGCAAGCTCAGGCCCGTCTGATGCTGTGGGAAGATGATGATCAGACAGC 3143

OY      355  LeuHISASPGInGlyHISerHisLeuGlyAspPro 366
          |||||||
DB      3144  CTTCATGACCAAGGCCACACCATCTGGGGGACCCC 3179

RESULT 10
LOCUS      AC074203          42479 bp    DNA      linear    PRI 13-OCT-2000
DEFINITION Homo sapiens chromosome 22q11 clone cos4, complete sequence.
ACCESSION  AC074203
VERSION     AC074203.3  GI:9625348
KEYWORDS
SOURCE      Homo sapiens.
            HFG.
            Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1
            (bases 1 to 42479)

```

AUTHORS Yang, L., Hu, P., Wu, H., Tapia-Paez, I. and Roe, B. A.
 TITLE Homo sapiens Chromosome 22q11 Cosmid cos4
 JOURNAL Unpublished
 REMARK This clone is a chimeric cosmid from a patient with t(11;22) translocation and it is from the derivative chromosome 11. There are sequences from both chromosome 11 and 22 in this cosmid 2 (bases 1 to 42479)

REFERENCE: Yang, L., Hu, P., Wu, H., Tapia-Paez, I. and Roe, B. A.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUL-2000) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 3 (bases 1 to 42479)
 Yang, L., Hu, P., Wu, H., Tapia-Paez, I. and Roe, B. A.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-2000) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 4 (bases 1 to 42479)

REFERENCE: Yang, L., Hu, P., Wu, H., Tapia-Paez, I. and Roe, B. A.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-2000) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA

COMMENT On Aug 1, 2000 this sequence version replaced g1:9441829.
 FEATURES
 source location/Qualifiers
 1. 42479
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="22q11"
 /clone="cos4"

HASK COUNT 11501 a 9314 c 9563 g 12101 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 length: 42479
 Score: 312.00 Matches: 312
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.25% Indels: 0
 DB: 9 Gaps: 0

US-09-842-364-3 (1-366) x AC074203 (1-42479)

QY 55 ThrLeuLysASPserLeuGlnGlnAspLeuAsnAsmMetLysPheLeuGlnLysLeu 74
 DB 33380 ACCCTGAAGAGACGCTTGAAGACCTCAACATATACAAAGTTCTTGGAAAGCTG 33439

QY 75 ArgProIeuSerGlySerGlnAlaProArgLeuProGlnAspProValClyMetArg 94
 DB 33440 AGGCTCTGAGTGGGAGCGAGCTCTCGGCTCCACAGACCGGTGGCATGGCGG 33499

QY 95 GlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 114
 DB 33500 CAGCTGACGAGAGAGTGGAGAGGCTGAGGCTCCCTCCACCTACATGGCAGCGC 33559

QY 115 HisGlnLeuValGlyTyrPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134
 DB 33560 CACGAGCTGGTGGTGGAAATTGGAGGCTTGGCGGCAACTGAAAGCCCTACACGATG 33619

QY 135 AspLeuMetGlnGlnValAlaLeuArgValGlnGlnLeuGlnGlnGlnGlnGlnGln 154
 DB 33620 GATCTGATGGAGCAGCTGCGCTGCGGTCAGAGCTGCAGAGCAGATTGGCGGTG 33679

QY 155 GlyGlnAspThrLysAlaGlnLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 174
 DB 33680 GCGGAGACACCAAGCCCACTGCTGCGGGGCGTGGACGAGCGGCTTGGCTGCGAG 33739

QY 175 GlyLeuGlnSerArgValAlaHisIsthrgLysArgPheLysGlnLeuPheHisPro 194
 DB 33740 GGACTGCAAGAGCGCGGTGGTGCACACCGCGCTTCAAGAGCTGTTCACCCATTC 33799

QY 195 AlaGlnSerLeuValSerGlyIleGlyArgHisValGlnGlnLeuHisArgSerValAla 214

DB 33800 GCCCAGACCTCGTGCAGCGGCATCGGCCGCCACCTGCAGGAGCTGCACCCACATCTGCT 33859

QY 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArg 234
 DB 33860 CCCAGCCCCCGCCAGCCCGCCCGCCCTGAGTGCCTGCGTGCAGAGCTCTCCCGGAG 33919

QY 235 LeuThrLeuLysAlaValAlaLeuHisAlaArgGlnGlnGlnGlnGlnGlnGlnGln 254
 DB 33920 CTCACGCTCAAGGCGCAGGCGCTTGCAGCGCATCGACAGAACTGGACCGAGCTGCG 33979

QY 255 GlnGlnLeuSerArgAlaPheAlaGlyThrGlyThrGlnGlnGlnGlnGlnGlnGlnGln 274
 DB 33980 GAAAGCTTCAGCAGAGCTTTCAGCGCATCGAGCAGTGAAGCGCGCCCGCCGAGCC 34039

QY 275 GlnMetLeuSerGlnGlnValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeu 294
 DB 34040 CACATGCTCTCCAGAGAGTGCAGCGCCACGACGCTTCCGCCAGCAGCCTTACCTG 34099

QY 295 GlnIleAlaAlaIleThrArgAlaIleAspGlnGlnGlnGlnGlnGlnGlnGlnGln 314
 DB 34100 CAGATAGCTGCTTCACTCGCGCATCGACAGAGACTGAGGAGTCCAGCAGAGCTG 34159

QY 315 AlaProProProProGlnHisSerAlaPheAlaProGlnGlnGlnGlnGlnGlnGlnGln 334
 DB 34160 GCGCCACCTCCACGAGCGCAGCTGCTTCCGCCAGAGCTTTCACAGACAGCAGCTGC 34219

QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGlnAspIleThrHisSer 354
 DB 34220 AAGCTTCTGAGCAAGCGCGAGCGCCCTGCTGATGACCTGTGGAGAACATCTCACAGC 34279

QY 355 LeuHisAspGlnGlnHisSerHisLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 366
 DB 34280 CTTCATGACCGAGGCGCAGCCATCTGGGGGAGCCCC 34315

RESULT 11
 AX063464
 LOCUS AX063464 81001 bp DNA linear PAT 24-JAN-2001
 DEFINITION Sequence 1 from Patent WO0100803.
 ACCESSION AX063464
 VERSION AX063464.1 GI:12541234
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Yeh, F., Denison, B., Bour, B., Bihain, B., Bougueleret, L., Duclert, A.
 and dumas mine Edwards, J.H.
 TITLE Apolipoprotein a-IV-related protein: polypeptide, polynucleotide
 sequences and diagnostic markers thereof
 JOURNAL Patent: WO 0100803-A1 04-JAN-2001;
 GENSET (FR)
 FEATURES
 location/Qualifiers
 1. 81001
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 primer_bind
 929..949
 /note="20-828..pu"
 primer_bind
 1220..1238
 /note="20-828-311.mis"
 misc_feature
 1227..1251
 /note="20-828-311.probe"
 variation
 1239..1259
 /note="20-828-311 : polymorphic base C or T"
 primer_bind
 1240..1258
 /note="20-828-311.mis complement"
 primer_bind
 1357..1377
 /note="20-828..rp complement"
 misc_feature
 10946..12946
 /note="5'regulatory region"
 primer_bind
 12029..12050
 /note="17-42..pu"

```

primer_bind      12328..12346
                  /note="17-42-319.mis"
misc_feature      12335..12359
                  /note="17-42-319.probe"
variation         12347
                  /note="17-42-319 : polymorphic base C or T"
primer_bind      12348..12366
                  /note="17-42-319.mis complement"
primer_bind      12581..12603
                  /note="17-42.rp complement"
exon              12947..12958
                  /note="exon 1"
exon              13470..13526
                  /note="exon 2"
exon              13641..13752
                  /note="exon 3"
exon              14271..15968
                  /note="exon 4"
primer_bind      14992..15012
                  /note="17-41.pu"
primer_bind      15222..15240
                  /note="17-41-250.mis"
misc_feature      15229..15253
                  /note="17-41-250.probe"
variation         15241
                  /note="17-41-250 : polymorphic base C or T"
primer_bind      15242..15260
                  /note="17-41-250.mis complement"
primer_bind      15460..15482
                  /note="17-41.rp complement"
misc_feature      15969..17969
                  /note="3' regulatory region"
primer_bind      42070..42090
                  /note="20-841.pu"
primer_bind      42199..42217
                  /note="20-841-149.mis"
misc_feature      42206..42230
                  /note="20-841-149.probe"
variation         42218
                  /note="20-841-149 : polymorphic base A or G"
primer_bind      42219..42237
                  /note="20-841-149.mis complement"
primer_bind      42572..42591
                  /note="20-841.rp complement"
primer_bind      45328..45347
                  /note="20-842.pu"
primer_bind      45423..45441
                  /note="20-842-115.mis"
misc_feature      45430..45454
                  /note="20-842-115.probe"
variation         45442
                  /note="20-842-115 : polymorphic base A or G"
primer_bind      45443..45461
                  /note="20-842-115.mis complement"
primer_bind      45863..45883
                  /note="20-842.rp complement"
primer_bind      76644..76664
                  /note="20-853.pu"
primer_bind      77039..77057
                  /note="20-853-415.mis"
misc_feature      77046..77070
                  /note="20-853-415.probe"
variation         77058
                  /note="20-853-415 : polymorphic base C or T"
primer_bind      77059..77077
                  /note="20-853-415.mis complement"
primer_bind      77166..77185
                  /note="20-853.rp complement"
BASE COUNT      20291 a 17705 c 19080 g 23919 t      6 others
ORIGIN
Alignment Scores: 0 Length: 81001
Pred. No.:

```

```

Score: 312.00 Matches: 312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.25% Indels: 0
DB: Gaps: 0

US-09-842-364-3 (1-366) x AX063464 (1-81001)

QY 55 ThrLeuLysAspSerLeuGlnAspLeuAsnMetAsnLysPheLeuGlnLysLeu 74
    |||
Db 14272 ACCCTGAAGAGAGAGCTTGTGAGCAGACCTCAACATATGAACAGTTCTGGAAAGCTG 14331

QY 75 ArgProLeuSerGlySerGlnAlaProArgLeuProGlnAspProValGlyMetArgArg 94
    |||
Db 14332 AGGCTCTGAGTGGAGGAGAGAGCTCTGCTCCACAGAGAGCCGCTGGGATGGGCGG 14391

QY 95 GlnLeuGlnGluGluLeuGlnGluValLysAlaArgLeuGlnProTyrMetAlaGlnAla 114
    |||
Db 14392 CAGCTGCAGAGAGAGTGTGAGAGAGTGAAGCTCGCTCCAGCCCTACATGGCAGAGCGC 14451

QY 115 HisGluLeuValGlyTyrAsnLeuGlnGlyLeuArgGlnGlnLysProTyrThrMet 134
    |||
Db 14452 CACGAGCTGGTGGGCTGGAAATTTGGAGGGCTTGGCGCAGCAACTGAACCTTACACGATG 14511

QY 135 AspleuMetGluGlnValAlaLeuArgValGlnGluLeuGlnGlnLysArgValAla 154
    |||
Db 14512 GATCTGATGGAGAGAGTGGCGCTGGCGTGCAGAGAGTGCAGAGAGCATTTGGCGTGGTG 14571

QY 155 GlyIleuAspThrLysAlaGlnLeuGlnGlyValAlaAspGlnAlaTrrPalaLeuGln 174
    |||
Db 14572 GCGGAAAGACCAAGAGCCAGCTTCTCGCGGCGCGCAGCAAGCTTGGCTTCTCCAG 14631

-QY 175 GlyLeuGlnSerArgValAlaHisThrGlyArgPheLysGlnLeuPheHisProTyr 194
    |||
Db 14632 GGACTGCAGAGAGCCGCTGGTGTGCACACAGCCGCCCTTCAAGAGACTTTCACCCATAC 14691

Db 14692 GCCGAGAGCTGTGTGAGGAGGATCGCGCCAGCTGCAGAGAGCTGCACCGCATGTGGCT 14751

QY 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgCys 234
    |||
Db 14752 CCGCAGCCGCCCGCCAGGCCGCCGCGCTCACTCCCTCGTCCAGAGTCTCTCCCGCAG 14811

QY 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnAsnLeuAspGlnLeuArg 254
    |||
Db 14812 CTCACGCTCAAGGCCAGAGCCCTGTGCAGCAGCATCCAGAGAACCTTGAGCCAGCTGGC 14871

QY 255 GluGluLeuSerArgAlaPheAlaGlyThrGluGluGluAlaGlyTyrAspPro 274
    |||
Db 14872 GAAGAGCTCAGCAGAGCTTGTGCAGGCACTGGAGTGGAAAGGGCGCGCCGAGCCCC 14931

QY 275 GlnMetLeuSerGluGlnValAlaArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeu 294
    |||
Db 14932 CAGATGCTCTCCGAGAGAGTGGCCCGCAGGACTTTCAGGCTTTCGCGCAGAGACTACTCTG 14991

QY 295 GlnIleAlaIlePheThrArgAlaAlaLeuAspGlnGluThrGluGlnValGlnGlnLeu 314
    |||
Db 14992 CAGATAGCTGCTTCACTACCTGGCGCATCCAGCAGAGAGTGAAGAGTCCAGCAGAGCTG 15051

QY 315 AlaProProProProGlnHisSerAlaPheAlaPheAlaProGluPheGlnGlnHisAspSerCys 334
    |||
Db 15052 GCGCCACTCCACAGGCGCCACAGTGCCTTGGCCCGAGTTTCAACAAACGAGAGTGGC 15111

QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGluAspIleThrHisSer 354
    |||
Db 15112 AAGGTTCGAGCAAGCTCAGAGCCCGTGTGATGACCTGTGGAAAGCATGACACACAC 15171

QY 355 LeuHisAspGlnGlnLysSerHisLeuGlnLysPro 366
    |||
Db 15172 CTTGATGACAGAGGCGCAGAGCATCTGGGGAGCC 15207

RESULT 12

```

AX469876 AX469876 81001 bp DNA linear PAT 09-AUG-2002
 LOCUS Sequence 1 from Patent WO02053734.
 DEFINITION AX469876
 ACCESSION AX469876
 VERSION AX469876.1 GI:22205149
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Bouqueret, L., Duclet, A., Clusel, C., Dumas, M. E., Yen-Polin, F.,
 Denison, B., Bihain, B., Bour, B., Ebbers-Reed, D., and Satter-Cid, L.
 TITLE Methods and compositions for inhibiting neoplastic cell growth
 JOURNAL Patent: WO 02053734-A 1 11-JUL-2002;
 GENSET SA (FR)
 FEATURES
 source location/Qualifiers
 1..81001
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 primer_bind 929..949
 /note="20-828-pu"
 primer_bind 1220..1238
 /note="20-828-311.mis"
 misc_binding 1227..1251
 /bound_moiety="20-828-311.probe"
 variation 1239
 /note="20-828-311 : polymorphic base C or T"
 primer_bind 1240..1258
 /note="20-828-311.mis complement"
 primer_bind 1357..1377
 /note="20-828-rp complement"
 misc_feature 10946..12946
 /note="5'regulatory region"
 primer_bind 12029..12050
 /note="17-42-pu"
 primer_bind 12328..12346
 /note="17-42-319.mis"
 misc_binding 12335..12359
 /bound_moiety="17-42-319.probe"
 variation 12347
 /note="17-42-319 : polymorphic base C or T"
 primer_bind 12348..12366
 /note="17-42-319.mis complement"
 primer_bind 12581..12603
 /note="17-42-rp complement"
 exon 12947..12958
 /number=1
 exon 13470..13526
 /number=2
 exon 13641..13752
 /number=3
 exon 14271..15968
 /number=4
 primer_bind 14992..15012
 /note="17-41-pu"
 primer_bind 15222..15240
 /note="17-41-250.mis"
 misc_binding 15229..15253
 /bound_moiety="17-41-250.probe"
 variation 15241
 /note="17-41-250 : polymorphic base C or T"
 primer_bind 15242..15260
 /note="17-41-250.mis complement"
 primer_bind 15460..15482
 /note="17-41-rp complement"
 misc_feature 15969..17969
 /note="3'regulatory region"
 primer_bind 42070..42090
 /note="20-841-pu"
 primer_bind 42199..42217
 /note="20-841-149.mis"
 misc_binding 42206..42230

variation /bound_moiety="20-841-149.probe"
 42218
 /note="20-841-149 : polymorphic base A or G"
 primer_bind 42219..42237
 /note="20-841-149.mis complement"
 primer_bind 42572..42591
 /note="20-841-rp complement"
 primer_bind 45328..45347
 /note="20-842-pu"
 primer_bind 45423..45441
 /note="20-842-115.mis"
 misc_binding 45430..45454
 /bound_moiety="20-842-115.probe"
 variation 45442
 /note="20-842-115 : polymorphic base A or G"
 primer_bind 45443..45461
 /note="20-842-115.mis complement"
 primer_bind 45863..45883
 /note="20-842-rp complement"
 primer_bind 76644..76664
 /note="20-853-pu"
 primer_bind 77039..77057
 /note="20-853-415.mis"
 misc_binding 77046..77070
 /bound_moiety="20-853-415.probe"
 variation 77058
 /note="20-853-415 : polymorphic base C or T"
 primer_bind 77059..77077
 /note="20-853-415.mis complement"
 primer_bind 77166..77185
 /note="20-853-rp complement"
 BASE COUNT 20291 a 17705 c 19080 g 23919 t 6 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 81001
 Score: 312.00 Matches: 312
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.25% Gaps: 0
 Caps: 0
 US-09-842-364-3 (1-366) x AX469876 (1-81001)
 QY 55 ThlleuLysAspSerleuGluGlnAspLeuAsnAsnMetCAsnLysPheLeuGluLysleu 74
 |||||
 Db 14272 ACCCTGAAGACAGCCCTGAGCAGACCTCAACAAATATGCAACATTCGCAAAAGCTG 14331
 QY 75 ArgProLeuSerGlySerGluAlaProArgLeuProGlnAspProValGlyMetArgArg 94
 |||||
 Db 14332 AGCCCTCTCAGTGGAGCGGAGCGCTCCGCTCCACAGAGCCCGGCGCATGCGCGG 14391
 QY 95 GlnLeuGlnGluGluLeuGluGluValLysAlaArgLeuGlnProGlyMetAlaGluAla 114
 |||||
 Db 14392 CAGCTGCAGAGAGAGTTGGAGGAGTGTAAGGCTCGCTCCAGGCCCACTGGAGAGGGG 14451
 QY 115 HisGluLeuValGlyTyrPheAsnLeuGluGlyLeuArgGlnGlnLeuLysProGlyThrMet 134
 |||||
 Db 14452 CACGAGCTGCTGGCTGGGAATTTGGAGGCTTCGCCACCACTGAAGCCCTACACGATG 14511
 QY 135 AspLeuMetGluGlnValAlaLeuArgValGlnGluLeuGlnGluGlnLeuArgValAla 154
 |||||
 Db 14512 GATCTGATGAGCAGAGTGGCCCTGCCCTGCACAGAGCTGCAGAGCAGTGGCGGTG 14571
 QY 155 GlyGluAspThrLysAlaGlnLeuLeuGluGlyValAspGluAlaThrAlaLeuLeuGln 174
 |||||
 Db 14572 GGGGAAGACACCAAGCCAGTGTGGGGGGGTGAGCAGAGGCTTGGCTTGTGCAG 14631
 QY 175 GlyLeuGlnSerArgValAlaHisHisThrGlyArgPheLysGluLeuPheHisProGly 194
 |||||
 Db 14632 GGACTGCAGAGCGCGTGTGGACACACCGCGCTTCAAGAGCTTTCACACCATAC 14691
 QY 195 AlaGluSerLeuValSerGlyIleGlyArgHisValGlnGluLeuHisArgSerValAla 214


```

|||||
Db 14692 GCCGAGAGCTGTGACGGGATCGGGCCACGTGCAGAGCTGCACCGCAGTGTGGCT 14751
QY 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
Db 14752 CCGCAGCCGCCCGCAGCCCGCGCGCTCAGTGTGCTGCGAGGTGCTCTCCCGAAG 14811
QY 235 LeuThrLeuysAlaLysAlaLeuHisAlaArgLysGlnGlnAsnLeuAspGlnLeuArg 254
Db 14812 CTCACGCTCAAGGCGCCCTCCACCGCAGCATCCAGCAACCTGCGACCTGCGCC 14871
QY 255 GluGluLeuSerArgAlaPheAlaGlyThrGlyThrGluGluGlyAlaGlyProAspPro 274
Db 14872 GAGAGGCTCAGCAGAGCTTTGCGAGGCACTGGAGCTAGAGAGAGGGCGCGCGACCC 14931
QY 275 GlnMetLeuSerGluGluValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeu 294
Db 14932 CAGATGCTCTCCGAGGAGTGCGCCAGCAGCTTCCGCGCAGACACCTACTG 14991
QY 295 GlnIleAlaAlaPheThrArgAlaIleAspGlnGluThrGluGluValGlnGlnGlnLeu 314
Db 14992 CAGATGCTCTCTCAGCTCCCGCCATCCAGCAGAGCTGAGAGGTCCAGCAGAGCTG 15051
QY 315 AlaProProProGluGluHisSerAlaPheAlaProGluPheGlnGlnThrAspSerGly 334
Db 15052 GCGCAGCTCCAGCAGGCGCAGAGTGGCTTCCGCGCAGAGTTCAACAAGCAGAGTGGC 15111
QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGluAspPheThrHisSer 354
Db 15112 AAGTTCTGAGCAAGCTGCGAGCCGCTGATACCTGTGGAGAGCATCATCATCAGCAGC 15171
QY 355 LeuHisAspGlnGluHisSerHisLeuGlnYAspPro 366
Db 15172 CTTTCATACCAAGGCGCCACACCATCTGCGGAGCCCC 15207

RESULT 13
AP001480/c
LOCUS
DEFINITION
SEQUENCE, 18 unordered pieces.
ACCESSION
AP001480
VERSION
AP001480.2 GI:8117338
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens DNA, clone: CTC-227C10.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 115027)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Tolokli, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 115,027 genomic DNA of 11g
Published only in Database (2000)
2 (bases 1 to 115027)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Tolokli, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (17-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail: hattori@psc.riken.go.jp,
URL: http://hqp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)
On May 30, 2000 this sequence version replaced gt:7288164.

COMMENT
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hqp.gsc.riken.go.jp/
Contact: hattori@psc.riken.go.jp
----- Project Information
Center project name: HumDrift11
Center clone name: CTC-227C10
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads

```

```

Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 107728 bases at least Q40
Consensus quality: 110501 bases at least Q30
Consensus quality: 111810 bases at least Q20
Insert size: 113327; sum-of-connigs
Quality coverage: 6.70x in Q20 bases; sum-of-connigs

NOTE: This is a 'working draft' sequence. It currently consists of
18 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.

1 15681 contig of 15681 bp in length
15782 30275 contig of 14494 bp in length
30376 40528 contig of 10153 bp in length
40629 50205 contig of 9577 bp in length
50306 56617 contig of 9312 bp in length
59718 66750 contig of 7033 bp in length
66851 73847 contig of 6997 bp in length
73948 79644 contig of 5697 bp in length
79745 85653 contig of 5909 bp in length
85754 91241 contig of 5488 bp in length
91242 91341 contig of 100 bp in length
91342 96419 contig of 5078 bp in length
96420 96519 contig of 100 bp in length
96520 100644 contig of 4125 bp in length
100645 100744 contig of 100 bp in length
100745 102923 contig of 2179 bp in length
102924 103023 contig of 100 bp in length
103024 105950 contig of 2227 bp in length
105951 106050 contig of 100 bp in length
106051 108938 contig of 2888 bp in length
108939 109038 contig of 100 bp in length
109039 111577 contig of 2539 bp in length

Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 15681: contig of 15681 bp in length
* 15682 15781: gap of 100 bp
* 30276 30275: contig of 14494 bp in length
* 30376 30375: gap of 100 bp
* 40529 40528: contig of 10153 bp in length
* 40629 40628: gap of 100 bp
* 50206 50205: contig of 9577 bp in length
* 50306 50305: gap of 100 bp
* 50306 59617: contig of 9312 bp in length
* 59618 59717: gap of 100 bp
* 59718 66750: contig of 7033 bp in length
* 66751 66850: gap of 100 bp
* 66851 73847: contig of 6997 bp in length
* 73848 73947: gap of 100 bp
* 73948 79644: contig of 5697 bp in length
* 79645 79744: gap of 100 bp
* 79745 85653: contig of 5909 bp in length
* 85654 85753: gap of 100 bp
* 85754 91241: contig of 5488 bp in length
* 91242 91341: gap of 100 bp
* 91342 96419: contig of 5078 bp in length
* 96420 96519: gap of 100 bp
* 96520 100644: contig of 4125 bp in length
* 100645 100744: gap of 100 bp
* 100745 102923: contig of 2179 bp in length
* 102924 103023: gap of 100 bp
* 103024 105950: contig of 2227 bp in length
* 105951 106050: gap of 100 bp
* 106051 108938: contig of 2888 bp in length
* 108939 109038: gap of 100 bp
* 109039 111577: contig of 2539 bp in length

```

* 11578 11677: gap of 100 bp
 * 11678 113487: contig of 1810 bp in length
 * 113488 113587: gap of 100 bp
 * 113588 115027: contig of 1440 bp in length.
 Location/Qualifiers
 1. 115027
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="CTC-227C10"

misc_feature
 /note="assembly_fragment"
 15782..30275
 /note="assembly_fragment"
 30376..40528
 /note="assembly_fragment"
 40629..50205
 /note="assembly_fragment"
 50306..59617
 /note="assembly_fragment"
 59718..66750
 /note="assembly_fragment"
 66851..73847
 /note="assembly_fragment"
 73948..79644
 /note="assembly_fragment"
 79745..85653
 /note="assembly_fragment"
 85754..91241
 /note="assembly_fragment"
 91342..96419
 /note="assembly_fragment"
 96520..100644
 /note="assembly_fragment"
 100745..102923
 /note="assembly_fragment"
 103024..105950
 /note="assembly_fragment"
 106051..108938
 /note="assembly_fragment"
 109039..111577
 /note="assembly_fragment"
 111678..113487
 /note="assembly_fragment"
 113588..115027
 /note="assembly_fragment"
 misc_feature
 /note="assembly_fragment"
 30684 a 26526 c 26291 g 29826 t 1700 others
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 115027
 Score: 312.00 Matches: 312
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.25% Indels: 0
 DB: 2 Gaps: 0

US-09-842-364-3 (1-366) x AP001480 (1-115027)

OY 55 ThrLeuLysAspSerLeuGluGlnAspIleuAsnAsnMetAsnLysPheLeuGluLysLeu 74
 DB 53163 ACCCTGAAGACAGCTTGAGCAAGACCTCAACATATGACAAAGTTCTGGAAAAAGCTG 53104
 OY 75 ArgProLeuSerGlySerGluAlaProAlaArgLeuProGlnAspProValGlyMetArgArg 94
 DB 53103 AGGCCTCTAGTGGAGCAAGCTCTCGGCTCCACAGAGACCCGGTGGCATGGCGCG 53044
 OY 95 GlnLeuGlnGluGlnLeuGlnGluValLysAlaArgLeuGlnProTyrMetAlaGluAla 114
 DB 53043 CAGCTGACAGAGAGATTGGAGAGGTGAAGGCTCGGCTCCAGCCCTACATGGAGAGGGG 52984
 OY 115 HisGluLeuValGlyTyrPasnLeuGlnGlyLeuArgGlnGlnLeuLysProTyrThrMet 134

DB 52983 CACAGCTGGTGGGCTTGGATTGGAGGCTTGGGAGCACTGAAGCCCTTACAGCATG 52924
 OY 135 AspLeuMetGluGlnValAlaLeuArgValGlnLysLeuGlnGlnLeuArgValVal 154
 DB 52923 GATCTGATGGAGCAGCGTGGCCCTTCGCTCCAGACAGCTGCAGAGCACTTCCCTGCTG 52864
 OY 155 GlyClnAspThrLysAlaGlnLeuLeuGlyClyValAspClnAlaTrrPalLeuLeuGln 174
 DB 52863 GGGGAGACACCAAGGCCCTGCTGGGGGCTGGACAGAGCTTGGCTTCTCAG 52804
 OY 175 GlyLeuGlnSerArgValAlaHisThrGlyValArgPheLysGluLeuPheHisProTyr 194
 DB 52803 GCAGTGCAGACCCCGCTGCTGCACACACCCCGCTTCAACAGCTTTCACACCATAC 52744
 OY 195 AlaGlnSerLeuValSerGlyIleGlyArgHisValGlnLeuHisAspSerValAla 214
 DB 52743 CCCAGACCCCTGGTGGAGCGGCATCGGCGCCAGCTGCAGAGCTGCACCCGAGTGGCT 52684
 OY 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
 DB 52683 CCGCAGCCCGCCCGCACCCCGCCGCTCAGTGGCTGGCGAGTGTCTCTCCGGAG 52624
 OY 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgTllegInGlnAsnLeuAspGlnLeuArg 254
 DB 52623 CTCAGCTCAAGGCCAAGCCCGCTGCACCCAGCTTCACACCAACTTCGACACCTGCC 52564
 OY 255 GluGlnLeuSerArgAlaPheAlaGlyThrGlyThrGlnGlnLysAlaGlyProAspPro 274
 DB 52563 GAAGAGCTTCAGCAGACGCTTGGCAGCGCAGCTGAGGAAGGGCGCGCCGAGCC 52504
 OY 275 GlnMetLeuSerGluGlnValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeu 294
 DB 52503 CAGTGTCTTCCGAGGAGGTGGCCAGCGACTTGAAGCTTTCGCGAGACCTTACTG 52444
 OY 295 GlnIleAlaPheThrAlaGlnAlaLeuAspGlnGlnThrGlnGlnValGlnGlnLeu 314
 DB 52443 CACAACCTGCTCCACTCCCGCCATCCAGCAGAGAGCTGAGAGGTCCACACAGCTG 52384
 OY 315 AlaProProProProGlnHisSerAlaPheAlaProGlnPheGlnGlnThrAspSerGly 334
 DB 52383 GCCCAGCTCCACAGCCAGCCAGCTGCTCCGCCAGAGTTTCAACAAACAGCAGCTG 52324
 OY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuProGlnLysThrHisSer 354
 DB 52323 AAGGTTCTGAGCAAGCTGAGGCGCGTCTGATGATCTGTTGGAGAACATCACTACAGC 52264
 OY 355 LeuHisAspGlnGlnHisSerHisLeuGlyAspPro 366
 DB 52263 CTTCATCAGCCAGCCAGCCAGCCATTCGCGCACCC 52228

RESULT 14
 AP001481 175582 bp DNA linear HTG 30-MAY-2000
 LOCUS Homo sapiens chromosome 11 clone CTC-270C21 map 11q, WORKING DRAFT
 DEFINITION SEQUENCE, 29 unordered pieces.
 ACCESSION AP001481
 VERSION AP001481.2 GI:8117339
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens DNA, Clone:CTC-270C21.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 175582)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 175,582 genomic DNA of 11q
 2 (bases 1 to 175582)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 TITLE
 JOURNAL
 Submitted (17-MAR-2000) Masahira Hattori, The Institute of Physical


```

/misc_feature      /note="assembly_fragment"
                    85624..94708
/misc_feature      /note="assembly_fragment"
                    94809..104076
/misc_feature      /note="assembly_fragment"
                    104177..112629
/misc_feature      /note="assembly_fragment"
                    112730..119309
/misc_feature      /note="assembly_fragment"
                    119410..126696
/misc_feature      /note="assembly_fragment"
                    126797..132422
/misc_feature      /note="assembly_fragment"
                    132523..137920
/misc_feature      /note="assembly_fragment"
                    138021..142190
/misc_feature      /note="assembly_fragment"
                    142291..145283
/misc_feature      /note="assembly_fragment"
                    145384..148550
/misc_feature      /note="assembly_fragment"
                    148651..151425
/misc_feature      /note="assembly_fragment"
                    151526..155432
/misc_feature      /note="assembly_fragment"
                    155533..158018
/misc_feature      /note="assembly_fragment"
                    158119..160661
/misc_feature      /note="assembly_fragment"
                    160762..163242
/misc_feature      /note="assembly_fragment"
                    163343..165254
/misc_feature      /note="assembly_fragment"
                    165355..167624

```

Alignment Scores:

```

Pred. No.:      0      Length:      175582
Score:          312.00  Matches:      312
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      85.25%  Indels: 0
DH:              2      Gaps: 0

```

US-09-842-364-3 (1-366) x AP001481 (1-175582)

```

QY 55 ThrLeuLysASPSeSerLeucluglInAspLeuAsnAsnMetAsnLysPheLeucluglLysLeu 74
|||||
Db 106171 ACCCTGAAGAAGACACCCCTTCAGCAGACACCCCAATATCAACAGTTCCTCGAAACGCTG 106112
|||||
QY 75 ArgProIeuSerClySerClyuAlaProArgIeuProGlnAspProValClyMetArgArg 94
|||||
Db 106111 AGGCTCTGAGTGGGAGGAGGAGCTCTCGGCTCCACAGGACCCGCTGGGCTATGGGCGG 106052
|||||
QY 95 GlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 114
|||||
Db 106051 CAGCTGACAGAGAGATTGAGAGAGAGCTGAAAGCTCCCTCCACCCCTACATGACAGAGCGG 105992
|||||
QY 115 HisGlnLeuValGlyTrpAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134
|||||
Db 105991 CACGAGCTGGTGGCTGGAAATTTGAGAGGCTTGGCGGACAGCACTGAACCCCTACACGAGTG 105932
|||||
QY 135 AspLeuMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 154
|||||
Db 105931 GATCTGATGAGAGACAGTGGGCTCGGCGTGGAGAGCTGACAGAGACAGTGGGCTGGTG 105872
|||||
QY 155 GlyIuAspPheTrpLysAlaGlnLeuLeucluglGlyValAspGluAlaTrpAlaLeuLeuGln 174
|||||
Db 105871 GGGGAGAGACACCAAGCCCACTTGGTGGGCGCTGACAGAGAGCTTGGCTTGGCTGCAG 105812
|||||
QY 175 GlyLeuGlnSerArgValValHisIsthrGlyArgPheLysGlnLeuPheHisProTrp 194
|||||
Db 105811 GGACTGACAGAGCCGCTGGTGGTGCACACACGCGCCCTTCAAGAGACTCTTCCACCCATAC 105752

```

```

QY 195 AlaGlnSerLeuValSerClyIleGlyArgHisValGlnGlnLeuHisArgSerValAla 214
|||||
Db 105751 GCGGAGAGCTGTGAGGAGGAGCTGGGCGCCACAGTGCAGAGCTGCACCGAGTGGCT 105692
|||||
QY 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
|||||
Db 105691 CCCAGAGCCGCCGCCAGAGCCGCCGCCCTCACCTGCCCTGCAGTGCAGTGCCTCCCGAAG 105632
|||||
QY 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnAsnLeuAspGlnLeuArg 254
|||||
Db 105631 CTCAGGCTCAAGGCCAAGGCCCTTCGACGACGACATTCACAGAACCTGGACCACTGGCG 105572
|||||
QY 255 GlnGlnLeuSerArgAlaPheAlaGlyTrpGlyThrGlnGlnGlnGlnGlnGlnGlnGlnGln 274
|||||
Db 105571 GAAAGAGCTAGAGAGAGCTTTCAGAGAGCTGGAGACTGAGAGAGGGGCGGCCGAGACCC 105512
|||||
QY 275 GlnMetLeuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 294
|||||
Db 105511 CAGAGTCTCTCCAGAGAGCTGGCGCCACGCTTCAGAGCTTCCTCCCGACAGCACTAACCTG 105452
|||||
QY 295 GlnIleAlaAlaPheThrArgAlaIleAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 314
|||||
Db 105451 CAGATAGCTGCTTCTACTGCGCCATCCAGCAGAGACTGAGAGAGGCTCAGAGCAAGCTG 105392
|||||
QY 315 AlaProProProProGlyHisSerAlaPheAlaProGluPheGlnGlnThrAspSerCly 334
|||||
Db 105391 GCGCCACCTCCACAGGAGCAGAGCTGCTGCGCCAGAGTTCACAAACAGACAGTGGC 105332
|||||
QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGlnAspLysIleThrIstSer 354
|||||
Db 105331 AAGCTTCTCAGCAAGCTGCAGGCGCCGCTGCGAGCTCTGCGAAGACATCTCAGACAGC 105272
|||||
QY 355 LeuHisAspGlnGlyHisSerHisLeuGlnAspPro 366
|||||
Db 105271 CTTCAATGACAGGAGGAGGAGCACTGCTGGGAGACCC 105226
|||||

```

RESULT 15

AC007707/c

LOCUS

DEFINITION

AC007707

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

TITLE Direct Submission
 JOURNAL Submitted (25-NOV-1999) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 On Aug 25, 1999 this sequence version replaced gi:5757506.
 COMMENT Location/Qualifiers
 FEATURES
 source 1..188971
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q23"
 /clone="442e11 from RPC111 Library"
 /note="442e11 was originally named b1030"

BASE COUNT 53091 a 44581 c 42653 g 48646 t
 ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	188971
Score:	312.00	Matches:	312
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	85.25%	Indels:	0
DB:	9	Gaps:	0

US-09-842-364-3 (1-366) x ACC007707 (1-188971)

QY 55 ThrLeuLysAspSerLeuGlnGlnAspLysMetAsnLysPheLeuGlnLysLeu 74
 DB 76729 ACCCTGAAGACACCTTGAGACACCTTCACCAATATGACACAGTTCTTGGAAAGCTTG 76670
 QY 75 ArgProLeuSerGlySerGluAlaProArgLeuProGlnAspProValGlyMetArgArg 94
 DB 76669 AGCCCTGAGTGGAGGAGGAGGCTCGCTCCACAGACCCGGTGGCATGGCCGGC 76610
 QY 95 GlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 114
 DB 76609 CAGCTGACAGAGAGGTGGAGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGGAGGTGG 76550
 QY 115 HisGlnLeuValGlyTrpAsnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 134
 DB 76549 CACGAGCTGGTGGGCGAATTGGAGGCTTGGCGGACCACTGAAGCCCTACACAGATG 76490
 QY 135 AspLeuMetGlnGlnValAlaLeuArgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 154
 DB 76489 GATCTGATGACACAGCTGCTCCGCGCAGCAGCTCCACAGCAGCTTGGCGCTG 76430
 QY 155 GlyIleuAspThrLysAlaGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 174
 DB 76429 GGGGAAACACCCAGGCGGCTTGGGGGGGCTGGAGCAGAGCTTGGCTGGCAG 76370
 QY 175 GlyLeuGlnSerArgValAlaHisIleThrGlyArgPheLysGlnLeuPheHisProTyr 194
 DB 76369 GGACTGACAGACCGCGCTGCTCCACACACCGCCCTTCAAGACCTCTTCCACCCATAC 76310
 QY 195 AlaGlnSerLeuValSerGlyTleGlyArgHisValGlnGlnGlnGlnGlnGlnGlnGln 214
 DB 76309 GCCGAGAGCTGCTGAGCGCCATCGGGCGCCACCTGACAGCTGACCCAGTGGCT 76250
 QY 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
 DB 76249 CCGCAGGCGCCCGCAGCGCCCGCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76190
 QY 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGln 254
 DB 76189 CTCACGCTCAAGGCCAGGCGCTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 76130
 QY 255 GlnGlnLeuSerArgAlaPheAlaGlyThrGlyThrGlnGlnGlnGlnGlnGlnGlnGlnGln 274
 DB 76129 GAAGAGCTCAGCAGACCTTTCAGGCACTGGAGCTGAGAGAGGCGCCCGCCGAGCC 76070
 QY 275 GlnMetLeuSerGlnGlnValArgGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 294
 DB 76069 CAGATGCTCTCCGAGAGGTGGCGCAGCTTACAGGCTTCCGCGCAGGACACTTACCTG 76010

QY 295 GlnIleAlaAlaPheThrArgAlaIleAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 314
 DB 76009 CAGATGACCTGCTTCACTGCTCCCATCGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 75950
 QY 315 AlaProProProGlyHisSerAlaPheAlaProGlnPheGlnGlnGlnGlnGlnGlnGlnGln 334
 DB 75949 GCGCACCTCCACAGGCGCAGAGTGGCCGAGAGTTTCAACAAACAGACAGTGGC 75890
 QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGluAspLleThrHisSer 354
 DB 75889 AAGTCTGTGACAGCAGCTGAGGCGCTGATGACCTGTGGAGAGACATCATCATCAGC 75830
 QY 355 LeuHisAspGlnGlyHisSerHisLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 366
 DB 75829 CTTCATGACCCAGGCGCCACAGCCATCTGGGGACCC 75794

Search completed: January 23, 2003, 18:46:04
 Job time : 3456 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 23, 2003, 17:46:52 : Search time 2033 Seconds

(without alignments)
2915.666 Million cell updates/sec

Title: US-09-842-364-3

Perfect score: 366
Sequence: 1 MASMAAVLTWALALISAFSA.....LWEDITHSLHDOGHSHLGD 366

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 6

Total number of hits satisfying chosen parameters: 231638

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-xml
-O/cgcn2.1/USPTO.spool/US09842364/runal.16012003.164829.10661/app.query.fasta.1.519
-DB-PST -QFMT-fastap -SUFFIX-olig.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE-quality -THR_MIN=6 -ALIGN=15 -MODE-LOCAL -OUTFMT-ptc
-NORR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09842364 -CGCN.1.1.1349 -grunal.16012003.164829.10661 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAL -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the local score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
C	1	75	20.5	696	17	AG097513	AG097513 Pan trogl
C	2	26	7.1	524	9	A1265480	A1265480 u102e07.x
C	3	26	7.1	542	12	Bf659944	Bf659944 ma25e08.
C	4	26	7.1	549	13	B1148686	B1148686 602912117
C	5	26	7.1	554	9	A15288320	A15288320 u196a04.y
C	6	26	7.1	606	13	B1144716	B1144716 602909506
C	7	26	7.1	608	12	Bf384518	Bf384518 602046482
C	8	26	7.1	612	13	B1221610	B1221610 602936921
C	9	26	7.1	612	9	AA987093	AA987093 uc74f12.y
C	10	26	7.1	617	12	Bf384624	Bf384624 602046895
C	11	26	7.1	645	12	Bf234067	Bf234067 602024986
C	12	26	7.1	648	12	Bf385873	Bf385873 602046040
C	13	26	7.1	660	12	Bf385843	Bf385843 602046010
C	14	26	7.1	663	13	B1220419	B1220419 602935137
C	15	26	7.1	672	12	Bf384592	Bf384592 602046861
C	16	26	7.1	672	13	B1332079	B1332079 602981411
C	17	26	7.1	699	12	Bf385663	Bf385663 602047318
C	18	26	7.1	705	13	B1144468	B1144468 602908638
C	19	26	7.1	705	13	B1255671	B1255671 602979053
C	20	26	7.1	716	13	B1145333	B1145333 602910476
C	21	26	7.1	722	12	Bf384847	Bf384847 602046543
C	22	26	7.1	722	12	Bf385520	Bf385520 602047065
C	23	26	7.1	727	12	Bf383715	Bf383715 602044622
C	24	26	7.1	735	13	B1145765	B1145765 602909228
C	25	26	7.1	746	13	B1331565	B1331565 602908983
C	26	26	7.1	758	12	Bf237009	Bf237009 602026641
C	27	26	7.1	762	12	Bf233452	Bf233452 602024241
C	28	26	7.1	767	13	B1149103	B1149103 602911511
C	29	26	7.1	767	13	B1246578	B1246578 602958496
C	30	26	7.1	769	13	B1144954	B1144954 602909602
C	31	26	7.1	773	13	B1145650	B1145650 602910686
C	32	26	7.1	778	13	B1145877	B1145877 602911175
C	33	26	7.1	789	13	B1247725	B1247725 602959363
C	34	26	7.1	791	12	Bf233098	Bf233098 602023839
C	35	26	7.1	791	13	B1147941	B1147941 602912655
C	36	26	7.1	792	13	B1145009	B1145009 602909669
C	37	26	7.1	792	13	B1146191	B1146191 602911005
C	38	26	7.1	793	13	B1144673	B1144673 602909969
C	39	26	7.1	798	12	Bf234162	Bf234162 602026017
C	40	26	7.1	804	13	B1145831	B1145831 602911107
C	41	26	7.1	805	13	B1148123	B1148123 602912487
C	42	26	7.1	806	13	B1148302	B1148302 602914316
C	43	26	7.1	809	12	Bf233225	Bf233225 602023779
C	44	26	7.1	815	13	B1331154	B1331154 602984189
C	45	26	7.1	820	13	B1332727	B1332727 602984481

ALIGNMENTS

RESULT 1
AC097513/c
LOCUS AC097513
DEFINITION Pan troglodytes DNA, clone: PTB-099D09.F, genomic survey sequence.
ACCESSION AC097513
VERSION AC097513.1 GI:16718030
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphoblast DNA, clone: PTB-099D09.F.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totok, Y., Watanabe, H. and Sakaki, Y.

REFERENCE 1 (bases 1 to 542)
 AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Other-ESTs: maa25c08.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

FEATURES
 source MGI:1454190
 Seq primer: -40RP from GIBCO
 High quality sequence stop: 435.
 Location/Qualifiers
 1..542

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:3812078"
 /clone_id="NCI_CGAP_L1110"
 /sex="female"
 /dev_stage="10 weeks"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.6 kb. Library constructed by Life
 Technologies."

BASE COUNT 130 a 151 c 182 g 78 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 6,35e-14 Length: 542
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 12 Gaps: 0

US-09-842-364-3 (1-366) x BF659944 (1-542)

OY 144 ValGInGluLeuGlnGlnLeuArgValValGlyGluAspThrLysAlaGlnLeu 163
 |||||||
 DB 300 CTGCAGCAGCTGCAGCAGCAGCTCCCTGCGCCAGACAGACCAAGCCTCCTCG 359

OY 164 GlyGlyValAspGluAla 169
 |||||||
 DB 360 GGGGGCGTGGAGCAGGCG 377

RESULT 4
 B1148686 549 bp mRNA linear EST 05-JUL-2001
 LOCUS 60291211F1 NCI_CGAP_L19 Mus musculus CDNA clone IMAGE:5052965 5'
 DEFINITION mRNA sequence.

ACCESSION B1148686
 VERSION B1148686.1 GI:114608687
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 549)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1144 row: b column: 06
 High quality sequence stop: 549.
 Location/Qualifiers
 1..549

FEATURES
 source /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5052965"
 /clone_id="NCI_CGAP_L19"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP library."
 BASE COUNT 133 a 145 c 189 g 82 t
 ORIGIN

Alignment Scores:
 Pred. No.: 6.43e-14 Length: 549
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 13 Gaps: 0

US-09-842-364-3 (1-366) x B1148686 (1-549)

OY 144 ValGInGluLeuGlnGlnLeuArgValValGlyGluAspThrLysAlaGlnLeu 163
 |||||||
 DB 432 CTGCAGCAGCTGCAGCAGCAGCTCCCTGCGCCAGACAGACCAAGCCTCCTCG 491

OY 164 GlyGlyValAspGluAla 169
 |||||||
 DB 492 GGGGGCGTGGAGCAGGCG 509

RESULT 5
 A1528320 554 bp mRNA linear EST 18-MAR-1999
 LOCUS u196a04.y1 Sugano mouse liver mla Mus musculus CDNA clone
 IMAGE:1890222 5' similar to SW:AVP4_MACPA_P33621 APOLIPROTEIN
 A-IV PRECURSOR ; mRNA sequence.

ACCESSION A1528320
 VERSION A1528320.1 GI:4442455
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 554)
 AUTHORS Mairra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person
 B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Riltter
 E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Mairra M/washu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu

COMMENT This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974546
 Seq primer: custom primer used
 High quality sequence stop: 513.

Alignment Scores:

Pred. No.: 7.05e-14 Length: 608
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 12 Caps: 0

US-09-842-364-3 (1-366) x BF384518 (1-608)

OY 144 ValGlnGluLeuGlnGluGlnLeuArgValValGlyGlnAspThrLysAlaGlnLeu 163

DB 424 CTGCAGAGCTGCAGACACACCTCCCTGCGAGACAGACCAAGCTCAGCTCTG 483

OY 164 GlyGlyValAspGluAla 169

DB 484 GGGGGCGTGGAGAGCGC 501

RESULT 8

LOCUS B1221610 608 bp mRNA linear EST 11-JUL-2001

DEFINITION 602936921F1 NCL_CGAP_L19 Mus musculus cDNA clone IMAGE:5100364 5',

ACCESSION B1221610 mRNA sequence.

VERSION B1221610.1 GI:14675054

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapsof@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: L1M11241 row: 1 column: 05

High quality sequence stop: 608.

Location/Qualifiers

1..608

/organism="Mus musculus"

/strain="FVH/N"

/db_xref="taxon:10090"

/clone="IMAGE:5100364"

/clone_1lb="NCI CGAP_L19"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: Liver; Vector: PCMV-SPORE6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.9 kb. Constructed by Life

Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 149 a 162 c 203 g 94 t

ORIGIN

Alignment Scores:

Pred. No.: 7.05e-14 Length: 608

Score: 26.00 Matches: 26

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 7.10% Indels: 0

DB: 13 Caps: 0

US-09-842-364-3 (1-366) x B1221610 (1-608)

OY 144 ValGlnGluLeuGlnGluGlnLeuArgValValGlyGlnAspThrLysAlaGlnLeu 163

DB 435 CTGCAGAGCTGCAGACACACCTCCCTGCGAGACAGACCAAGCTCAGCTCTG 494

OY 164 GlyGlyValAspGluAla 169

DB 495 GGGGGCGTGGAGAGCGC 512

RESULT 9

LOCUS AA987093

DEFINITION AA987093 612 bp mRNA linear EST 28-MAY-1998

uc74112.y1 Sugano mouse liver mlia Mus musculus cDNA clone

IMAGE:1431407 5' similar to SW:AP4_MACPA P33621 APOLOPROTEIN

A-IV PRECURSOR ; mRNA sequence.

ACCESSION AA987093

VERSION AA987093.1 GI:3167856

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE Matra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The Washu-HMI Mouse EST Project

Unpublished (1996)

Contact: Matra M/Mouse EST Project

Washu-HMI Mouse EST Project

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:915475

Seq primer: primer name ambiguous

High quality sequence stop: 508.

Location/Qualifiers

1..612

/organism="Mus musculus"

/strain="C57BL"

/db_xref="taxon:10090"

/clone="IMAGE:1431407"

/clone_1lb="Sugano mouse liver mlia"

/sex="female"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: liver; Vector: pME18S-Fl3; Site_1: DraIII

(CACCTGTCG); Site_2: DraIII (CACCTGTCG); 1st strand cDNA

was primed with an oligo(dT) primer

[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was

ligated to a DraIII adaptor (TGTGGCTCTGTCG), digested

and cloned into distinct DraIII sites of the pME18S-Fl3

vector (5' site CACCTGTCG, 3' site CACCTGTCG). XhoI should

be used to isolate the cDNA insert. Size selection was

performed to exclude fragments <1.5kb. Library

constructed by Dr. Sumio Sugano (University of Tokyo

Institute of Medical Science). Custom primers for

sequencing: 5' end primer CTTCGCTCTTAAAGCTGCG and 3' end

primer CGACCTGACCTGACGACAC.

BASE COUNT 152 a 166 c 202 g 92 t

ORIGIN

Alignment Scores:

Pred. No.: 7.1e-14 Length: 612

Score: 26.00 Matches: 26

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 7.10% Indels: 0

DB: 9 Caps: 0

US-09-842-364-3 (1-366) x AA987093 (1-612)

```

QY 144 ValGIngluLeuGIngluGInleuAryValGlyIuAspThrIlysaIaGInleu 163
DB 443 GTGCGAGAGCTGCAGACACCTGCGCTGCTGGAGAAACACCAAGCTCAGCTCTG 502
QY 164 GtGtYtYtAAspGtYtA 169
DB 503 GGGGGCGTGCAGAGCGG 520

RESULT 10
BF384624 617 bp mRNA linear EST 27-NOV-2000
LOCUS 602046895F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4196542 5',
DEFINITION mRNA sequence.
ACCESSION BF384624
VERSION BF384624
KEYWORDS GI:11365929
SOURCE EST.
ORGANISM house mouse.
            Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 617)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM9533 row: m column: 23
            High quality sequence stop: 615.
            Location/Qualifiers
                1. 617
                /organism="Mus musculus"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone_image="4196542"
                /clone_lib="NCI_CGAP_L19"
                /lab_host="DH10B (TI phage-resistant)"
                /note="Organ: liver; Vector: pCMV-Sport6; Site: 1: NotI;
                Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.9 kb. Constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 152 a 164 c 206 g 95 t
ORIGIN

Alignment Scores:
Pred. No.: 7,15e-14 Length: 617
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.10% Indels: 0
DB: 12 Gaps: 0

US-09-842-364-3 (1-366) x BF384624 (1-617)
QY 144 ValGIngluLeuGIngluGInleuAryValGlyIuAspThrIlysaIaGInleu 163
DB 423 GTGCGAGAGCTGCAGACACCTGCGCTGCTGGAGAAACACCAAGCTCAGCTCTG 482
QY 164 GtGtYtYtAAspGtYtA 169
DB 483 GGGGGCGTGCAGAGCGG 500

RESULT 11
BF234067 645 bp mRNA linear EST 14-NOV-2000
LOCUS 602024986F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4160397 5',
DEFINITION

```

```

ACCESSION BF234067
VERSION BF234067.1 GI:11145152
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 645)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM9439 row: k column: 22
            High quality sequence stop: 639.
            Location/Qualifiers
                1. 645
                /organism="Mus musculus"
                /strain="FVB/N"
                /db_xref="taxon:10090"
                /clone_image="4160397"
                /clone_lib="NCI_CGAP_L19"
                /lab_host="DH10B (TI phage-resistant)"
                /note="Organ: liver; Vector: pCMV-Sport6; Site: 1: NotI;
                Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.9 kb. Constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 153 a 180 c 212 g 100 t
ORIGIN

Alignment Scores:
Pred. No.: 7,44e-14 Length: 645
Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.10% Indels: 0
DB: 12 Gaps: 0

US-09-842-364-3 (1-366) x BF234067 (1-645)
QY 144 ValGIngluLeuGIngluGInleuAryValGlyIuAspThrIlysaIaGInleu 163
DB 413 GTGCGAGAGCTGCAGACACCTGCGCTGCTGGAGAAACACCAAGCTCAGCTCTG 472
QY 164 GtGtYtYtAAspGtYtA 169
DB 473 GGGGGCGTGCAGAGCGG 490

RESULT 12
BF385873 648 bp mRNA linear EST 27-NOV-2000
LOCUS 602046040F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4195808 5',
DEFINITION mRNA sequence.
ACCESSION BF385873
VERSION BF385873.1 GI:11367178
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 648)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

```

Email: c9apbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM9531 row: 0 column: 09
 High quality sequence stop: 646.
 Location/Qualifiers

/strain="FVB/N"
 /db.xref="taxon:10090"
 /clone="IMAGE:4195514"
 /clone.lib="NCI_CGAP_L19"
 /lab.host="DH10B (T1 phage-resistant)"
 /note="Organ: Liver; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo df. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."
 BASE COUNT 157 a 185 c 217 g 101 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7.48e-14 Length: 648
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 12 Gaps: 0
 US-09-842-364-3 (1-366) x BF385843 (1-648)
 BASE COUNT 155 a 177 c 215 g 101 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7.6e-14 Length: 660
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 12 Gaps: 0
 US-09-842-364-3 (1-366) x BF385843 (1-660)
 ORIGIN

Alignment Scores:
 Pred. No.: 7.48e-14 Length: 648
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 12 Gaps: 0
 US-09-842-364-3 (1-366) x BF385873 (1-648)
 ORIGIN

Alignment Scores:
 Pred. No.: 7.6e-14 Length: 660
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 12 Gaps: 0
 US-09-842-364-3 (1-366) x BF385843 (1-660)
 ORIGIN

US-09-842-364-3 (1-366) x BF385873 (1-648)
 ORIGIN

US-09-842-364-3 (1-366) x BF385843 (1-660)
 ORIGIN

FEATURES
 SOURCE
 1. 660
 /organism="Mus musculus"
 /strain="FVB/N"
 /db.xref="taxon:10090"
 /clone="IMAGE:5098473"
 /clone.lib="NCI_CGAP_L19"
 /lab.host="DH10B (T1 phage-resistant)"
 /note="Organ: Liver; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo df. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."
 BASE COUNT 163 a 183 c 216 g 101 t
 ORIGIN

FEATURES
 SOURCE
 1. 663
 /organism="Mus musculus"
 /strain="FVB/N"
 /db.xref="taxon:10090"
 /clone="IMAGE:5098473"
 /clone.lib="NCI_CGAP_L19"
 /lab.host="DH10B (T1 phage-resistant)"
 /note="Organ: Liver; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo df. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."
 BASE COUNT 163 a 183 c 216 g 101 t
 ORIGIN

Pred. No.: 7.63e-14 Length: 663
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 13 Gaps: 0

US-09-842-364-3 (1-366) x B1220419 (1-663)

OY 144 ValGInCluEnuGInCluGInLeuArGValValGlyCluAspThrLysAlaGInLeu 163
 Db 425 GTGAGAGAGCTGCAGAGAGCTGCCTGCTGTGGAGAGACACCAAGCTCAGCTCCTG 484
 OY 164 GlyGlyValAspGluAla 169
 Db 485 GCGGGCTGCAGCAGCGC 502

RESULT 15

BF384592 672 bp mRNA linear EST 27-NOV-2000
 LOCUS 602046861F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4196438 5',
 DEFINITION mRNA sequence.

ACCESSION BF384592 GI:11365897

VERSION BF384592

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE
 AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE NIH-MGC http://mgc.nci.nih.gov/ Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps@email.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.C.E. Consortium/LINL at:
 http://image.lnl.gov
 Plate: L1AM9533 row: 1 column: 15
 High quality sequence stop: 672.
 Location/Qualifiers

FEATURES

source

1. 672
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:4196438"
 /clone_lib="NCI_CGAP_Li9"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: liver; Vector: pCMV-Sport6; Site_1: NCI;
 Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.9 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 153 a 209 c 203 g 107 t
 ORIGIN

Alignment Scores:

Pred. No.: 7.73e-14 Length: 672
 Score: 26.00 Matches: 26
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.10% Indels: 0
 DB: 12 Gaps: 0

US-09-842-364-3 (1-366) x BF384592 (1-672)

OY 144 ValGInCluEnuGInCluGInLeuArGValValGlyCluAspThrLysAlaGInLeu 163
 Db 53 GTGAGAGAGCTGCAGAGAGCTGCCTGCTGTGGAGAGACACCAAGCTCAGCTCCTG 112

OY 164 GlyGlyValAspGluAla 169
 Db 113 GCGGGCTGCAGCAGCGC 130
 Search completed: January 23, 2003, 19:22:41
 Job time : 2037 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 23, 2003, 17:46:02 ; Search time 262 Seconds

(without alignments)
3145.921 Million cell updates/sec

Title: US-09-842-364-3

Perfect score: 366
Sequence: 1 MASMAVLVWALALJSAFSA.....LMEDITHSLHDGSHLGD 366

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delcext 7.0

Searched: 2185239 segs, 112599159 residues

Word size: 6

Total number of hits satisfying chosen parameters: 29428

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n model -DEV-xlh
-O-Cgrr2_L/USPRO.spool/US09842364/runat.16012003.164829.10640/app_query.fasta.1.519
-DB-N/Geneseq.101002 -OFRMT-fastap -SUFFIX-olig.rng -MIMATCH-0.1 -LOOPEL-0
-LOOPEXT-0 -UNIT-5-bits -THR-SCORE-quality -THR_MIN-6 -ALIGN-15 -MODE-LOCAL
-LIST-45 -DOCALLIGN-200 -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USRR-US09842364.@CGN.1.1.187 @runat.16012003.164829.10640 -NCPU-6 -ICPU-3
-NO_XLPHY -NO_MMAP -LARGOUDRY -NEG_SCORES-0 -WAIT -LONGLOG -DEV_TIMEOUT-120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP-60 -XGAPEXT-60 -Fgapop-6 -Fgapext-7
-YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database :
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	366	100.0	1322	21	AAZ45874	Nucleotide sequenc
2	366	100.0	1858	22	AAJ19216	Human CG122 (or C8
3	366	100.0	1859	22	AA522859	Human cDNA encodin
4	366	100.0	1867	21	AAZ45875	Nucleotide sequenc
5	366	100.0	1877	22	AAJ19236	Human lipid metabo
6	366	100.0	1877	22	AA522623	Human cDNA encodin
7	366	100.0	1879	22	AA520036	Human apolipoprote
8	366	100.0	1894	21	AA87740	Human secreted pro
9	366	100.0	1894	22	AA874022	CDNA encoding huma
10	312	85.2	2249	22	AA89395	Human digestive sy
11	312	85.2	2250	22	AA89396	Human digestive sy
12	312	85.2	2250	22	AA89397	Human digestive sy
13	312	85.2	5381	22	AA89397	Apolipoprotein A-1
14	312	85.2	81001	22	AA89397	Human apolipoprote
15	279	76.2	1371	24	ABK40279	CDNA encoding huma
16	279	76.2	1371	24	ABK40279	CDNA encoding huma
17	279	76.2	1393	22	AA89397	Human secreted pro
18	213	58.2	1334	22	AA85065	Atherosclerosis-as
19	176	48.1	592	22	AA85065	Receptor #93 parti
20	70	19.1	468	22	AA89396	Human secreted pro
21	68	18.6	2049	23	AA573514	Human encoding novel
22	33	9.0	331	22	AA89396	Human secreted pro
23	33	9.0	538	22	AA89396	Human digestive sy
24	16	4.4	1282	21	AA245872	Nucleotide sequenc
25	16	4.4	1834	21	AA245873	Nucleotide sequenc
26	16	4.4	1834	21	AA245873	Rat RAP3 gene. Ka
27	14	3.8	831	23	AA573513	DNA encoding novel
28	11	3.0	5377	24	ABK28379	DNA transcription
29	9	2.5	1376	21	AA89396	Arabidopsis thalia
30	9	2.5	1743	24	AA520017	Partial cDNA encod
31	9	2.5	2850	23	AA53767	Helicobacter pylor
32	8	2.2	5377	24	ABK28380	Unit of repeated s
33	8	2.2	48	15	AA057654	Human spliced tran
34	8	2.2	60	24	ABN32618	Human dopamine D4
35	8	2.2	105	16	AA740295	Non-overlapping re
36	8	2.2	105	20	AAV71234	Human dopamine D4
37	8	2.2	201	16	AA740296	Non-overlapping re
38	8	2.2	201	20	AAV71235	Human dopamine D4
39	8	2.2	279	22	AA503067	Non-overlapping re
40	8	2.2	286	21	AA87100	Human diagnostic a
41	8	2.2	345	16	AA740297	Rat hepatocyte car
42	8	2.2	345	20	AAV71236	Human dopamine D4
43	8	2.2	346	21	AA89396	Human secreted pro
44	8	2.2	397	19	AA660031	Nucleic acid G127
45	8	2.2	431	23	AA568995	DNA encoding novel

ALIGNMENTS

RESULT 1
AAZ45874 standard: cDNA: 1322 BP.

AAZ45874:
25-APR-2000 (first entry)
Nucleotide sequence of the 1322 bp cDNA sequence of human RAP3 gene.
RAP3; regeneration association protein 3; liver regeneration;
liver proliferation; ss.
Homo sapiens.

Key Location/Qualifiers
CDS 19..1110
FT /*tag- a

/product- "RAP3 protein"

FT
 XX
 PN WO200003013-A2.
 XX
 PD 20-JAN-2000.
 XX
 PF 12-JUL-1999; 99MO-EP04938.
 XX
 PR 10-JUL-1998; 98EP-0202336.
 XX
 PA (AMST-) AMSTERDAM MOLECULAR THERAPEUTICS BV.
 PI Chamileau RAFM, Groenink M, Van Der Vliet HN, Leegwater ACJ;
 XX WPI: 2000-147615/13.
 DR P-PSDB: AAY34562.
 XX
 PT Isolated RAP3 gene, protein and antibody useful for diagnosing liver
 PT regeneration and/or cell proliferation -
 XX
 PS Claim 3: Fig 6A: 42pp: English.

CC The present sequence encodes a human RAP3 (regeneration association
 CC protein 3). The RAP3 cDNA sequence was isolated from a human liver cDNA
 CC library. The RAP3 gene is involved in regeneration processes of the
 CC liver. The rat RAP3 gene was found to be upregulated 6 hours after
 CC partial hepatectomy, after which it was downregulated. The RAP3 cDNA
 CC sequence is useful as a source for PCR primers and probes, which are
 CC useful for detecting nucleotide sequences in a source material. The
 CC RAP3 cDNA sequence is useful as a marker of liver proliferation. The
 CC RAP3 protein is useful for the diagnosis of liver regeneration and liver
 CC cell proliferation. RAP3 antibodies, PCR primers and probes are useful
 CC for detecting the occurrence of liver cell proliferation in a patient.
 CC The RAP3 protein is also useful for enhancing the growth of
 CC regeneration of liver tissue comprising treating the liver tissue
 CC such as extracorporeal or intracorporeal.

SQ Sequence 1322 BP; 267 A; 408 C; 427 G; 220 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 1322
 Score: 366.00 Matches: 366
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Gaps: 0

US-09-842-364-3 (1-366) x AAY45874 (1-1322)

QY 1 MetAlaSerMetAlaAlaValIcuThrTrpAlaIeuAlaIeuSerAlaIphSerAla 20
 Db 10 ATGGCAAGCATGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 69
 QY 21 ThrGlnAlaArgIysGlyPheTrpAspTyrPheSerGlnThrSerGlyAspLysGlyArg 40
 Db 70 ACCCAGGCACGGAAGACCTTCTGGCACTACTCAGCCAGACACCGGGCAAAAGCAGC 129
 QY 41 ValGlnGlnIleHisGlnGlnIleMetAlaArgGlnProAlaThrIleuLysAspSerIleu 60
 Db 130 GTGGAGAGATGCATCGCAGGAAGATGGCTGGCAGCCCGGAGCCCGAAAGACAGCCCTT 189
 QY 61 GlnGlnAspLeuAsnAsnMetAlaLysPheLysGlnLysLeuArgProLysSerGlySer 80
 Db 190 GAGCAAGACCTCAACATATGAAAGAGTCTCTGAAAAGCTGAGGCTCTGAGTGGAGC 249
 QY 81 GluAlaProArgLeuProGlnAspProValGlyMetArgArgGlnLeuGlnGlnIleu 100
 Db 250 GAGGCTCTCGCTCCACAGCAGACCGGTGGCATCGCGGCGAGCTGAGAGAGAGTTG 309
 QY 101 GlnGlnValLysAlaArgLeuGlnProTyrMetAlaGlnAlaHisGlnLeuValGlyTyr 120
 Db 310 GAGGAGGTGAAGGCTCGCTCCAGCCCTACATGACAGGCGCAGCAGCTGTGGCTGG 369

QY 121 AsnLeuGlnGlyLeuArgGlnGlnIleuLysProTyrThrMetAspLeuMetGlnIleuVal 140
 Db 370 AATTGGAGGCTTGGCGGACAGTGAAGCCCTACAGATGATCTGATGAGACAGCTG 429
 QY 141 AlaLeuArgValGlnGlnIleuGlnGlnIleuArgValValGlyGlnAspThrLysAla 160
 Db 430 GCCCTGGCGGCTGCAGAGACTTCAGAGACAGTTCGCGGTGGTGGGGAAGACCAAGGCC 489
 QY 161 GlnLeuLeuGlnGlyValAspGlnAlaTrpAlaLeuLeuGlnGlnIleuGlnSerArgVal 180
 Db 490 CAGTTGCTGGGGGCGGTGACAGAGCTTGGCTTGTGTCAGAGGAGCTCAGAGCGCGCTG 549
 QY 181 ValHisHisThrGlnArgPheLysGlnLeuPheHisProTyrAlaGlnIleuValSer 200
 Db 550 GTGCACACACAGCGCGCTTCAAGAGCTTTCACCCATACAGCGGAGAGCTGATGAGC 609
 QY 201 GlyIleGlyArgHisValGlnGlnIleuHisArgSerValAlaProHisAlaProAlaSer 220
 Db 610 GGCATCGGCGGCCACAGCTGCAGAGACTGCACCCAGTGTGCTCCGACCCCGCCAGC 669
 QY 221 ProAlaArgLeuSerArgCysValGlnValLeuSerArgLysLeuThrLysAlaLys 240
 Db 670 CCGCGCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
 QY 241 AlaLeuHisAlaArgIleGlnGlnIleuAsnLeuAspGlnIleuArgGlnIleuValSer 260
 Db 730 GCGCTGCACGACGACCATCCACAGACCTGACACCTGACCTGACCTGACCTGACCTGAC 789
 QY 261 PheAlaGlyThrGlyThrGlnGlnGlnValGlyProAspProGlnMetLeuSerGlnGlu 280
 Db 790 TTTCAGGACCATGGGACGACGAGAAAGGCGCGCCGCGACCCCGCATGCTCTCCAGAG 849
 QY 281 ValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeuGlnIleAlaAlaPheThr 300
 Db 850 GTGGCTAGCCACTTCAGGCTTCTCCGCGAGACACTCTGACGATAGCTGCTGCTGCTG 909
 QY 301 ArgAlaIleAspGlnGlnIleuThrGlnGlnValGlnGlnIleuAlaProProGlnGly 320
 Db 910 CGCGGCTATCGACAGAGACTGAGAGAGTCCAGAGGAGCTGGGCGCCACTCCACAGGC 969
 QY 321 HisSerAlaPheAlaProGlnPheGlnGlnIleuThrAspSerGlyLysValLeuSerLys 340
 Db 970 CACAGTGGCTTGGCGCCGAGACTTTCACAAACAGACAGTGGCAAGCTTCTGAGCAAGCTG 1029
 QY 341 GlnAlaArgLeuAspAspLeuTrpGlnAspIleThrHisSerLeuHisAspGlnGlyHis 360
 Db 1030 CAGGCGGCTGTGATGACTGTGGGAAGCATCACTCAGACCTTCATGACAGGCGCAC 1089
 QY 361 SerHisLeuGlnGlyAspPro 366
 Db 1090 AGCCATCTGGGGGAGCCCT 1107

RESULT 2
 AAD19216
 ID AAD19216 standard; DNA; 1858 BP.
 XX
 AC AAD19216:
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human CG122 (or C868) DNA.
 XX
 KW Human: apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
 KW cardiovascular disease; lipid metabolism; myocardial infarction;
 KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KW neuroprotectant; cerebroprotective; ds.

OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 46..1146

```

FT      /tag= a          /*product= "Human CG122 (or C868) protein"
FT      sig_peptide    46..114
FT      /tag= b
FT      mal_peptide    115..1143
FT      /tag= c
FT      /product= "Human mature CG122 (or C868) protein"
XX
XX
XX
PD      WO200179446-A2.
XX
XX
PP      25-OCT-2001.
XX
XX
PP      16-APR-2001; 2001WO-USJ2529.
XX
XX
PR      14-APR-2000; 2000US-197137P.
PR      20-JUN-2000; 2000US-0598042.
PR      03-AUG-2000; 2000US-0631451.
PR      22-SEP-2000; 2000US-0667296.
PR      17-NOV-2000; 2000US-0714936.
XX
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI      Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ken F, Qian XB;
PI      Wang D;
DR      WPI: 2001-611724/70.
DR      P-PSDB: AAEL1922.
XX
XX
PT      Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein'
PT      receptor polypeptides, useful for preventing diagnosing and treating
PT      lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
XX
PS      Claim 1; Page 146-148; 266pp; English.
XX
XX
CC      The invention relates to polynucleotides encoding proteins CG122, CG179
CC      CG95, CG121, CG162, CG37 CG153 and CG168 which are related to proteins
CC      involved in lipid metabolism and cardiovascular disease such as human
CC      apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC      and protein sequences are useful for treating or preventing disorders
CC      associated with apolipoproteins, lipases and lipoprotein receptor (ALRP
CC      expression and for treating lipid metabolism, cardiovascular diseases
CC      and thrombosis. Antibodies against these proteins are useful for
CC      determining the presence of or predisposition to a disease associated
CC      with altered levels of these sequences. ALRP polypeptides are also
CC      useful for identifying agents (agonists and antagonists) that bind to
CC      them and cells expressing ALRP proteins are useful for identifying a
CC      therapeutic agent for use in treatment of a pathology related to
CC      aberrant expression or physiological interactions of this polypeptide.
CC      Vectors comprising these DNA and protein sequences are also useful for
CC      producing ALRP proteins. The nucleic acids and polypeptides of the
CC      invention are also useful for the treatment of occlusive cardiovascular
CC      diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC      thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC      or intracardiac thrombosis and stroke. The nucleotides of the invention
CC      are used in gene therapy. The present sequence is human CG122 (or C868)
CC      DNA.
XX
XX
SQ      Sequence 1858 BP; 379 A; 549 C; 582 G; 347 T; 1 other:
XX
XX
Alignment Scores:
Pred. No.: 0 Length: 1858
Score: 366.00 Matches: 366
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 22 Gaps: 0
US-09-842-364-3 (1-366) x AAD19216 (1-1858)
QY      1 MetlaasrMetAlaIvalleuThrTPralaleuaLaleuSeRaAlapheSeRaLa 20
        |||||||
DB      46 ATGGCAACATAGGCGCCTGCCTCACCTCGGGCTCTGTGCTCTTCCTTCACGCCATTTCGGCC 105

```


Db 919 TTGGACGACCTGGAGTGAAGGGGCGCGCCGACCCAGATGCTCTCGAGAG 978
 Qy 281 ValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeuGlnIleAlaAlaPheThr 300
 Db 979 GTGGCGCAGGACTTCAGGCTTCCGCGCAGACACCTGACATGAGTGTGCTTCACT 1038
 Qy 301 ArgAlaIleAspGlnIleThrGluGluValGlnGlnIleuAlaProProProGly 320
 Db 1039 CCGCCCATCGACACGACGACTGACAGCTCCACGAGCTGGCCGACCTCCACCGCC 1098
 Qy 321 HisSerAlaPheAlaProGluPheGlnGlnThrAspSerGlyValLeuSerLysLeu 340
 Db 1099 CACAGTGCCTTGGCCCGCAGAGTTTCAACAAACAGACAGTGGCAAGTTCGAGCAACTG 1158
 Qy 341 GlnAlaArgLeuAspAspLeuTrpGluAspIleThrHisSerLeuHisAspGlnGlyHis 360
 Db 1159 CAGGCCCGCTGATGATACCTGTGGAAACATCATCTACAGCTTCAATGACGAGGCCAC 1218
 Qy 361 SerHisLeuGlyAspPro 366
 Db 1219 ACCCATCTGGGGGACCC 1236
 RESULT 6
 AAS22623
 ID AAS22623 standard; cDNA; 1877 BP.
 AC AAS22623;
 XX AAS22623;
 DT 24-OCT-2001 (first entry)
 XX
 DE Human cDNA encoding a novel human protein #189.
 XX
 KM Human: novel protein; ss; Antinaemic; osteopathic; anti-inflammatory;
 KM immunomodulatory; cytoskeletal; neuroprotective; vulnery; nocotropic;
 KM anticonvulsant; antiallergic; cerebroprotective; antitumoral; antiviral;
 KM antidiabetic; antiallergic; dermatological; haemostatic; antiasthmatic;
 KM thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KM Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KM tissue regeneration; immune disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200155437-A2.
 XX
 PD 02-AUG-2001.
 XX
 PE 25-JAN-2001; 2001WO-US02623.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Dermanac RT;
 XX
 DR WPI: 2001-451939/48.
 DR P-PSDB; AAU14318.
 XX
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage.
 XX
 PS Claim 1; Page 429-431; 894pp; English.
 XX
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or

CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicits an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence encodes a protein of the invention.
 XX
 SQ Sequence 1877 BP; 409 A; 560 C; 584 G; 324 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 1877
 Score: 366.00 Matches: 366
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-842-364-3 (1-366) x AAS22623 (1-1877)
 Qy 1 MetaLaserMetaAlaValIleuThrTrpAlaLeuAlaLeuSerAlaPheSerAla 20
 Db 139 ATGGCAAGCATGGCTGGCGTCTCCTGCTGCTTCTTCAAGCTTTTCGGCC 198
 Qy 21 ThrGlnAlaArgLysGlyPheTrpAspTyrPheSerGlnThrSerGlyAspLysGlyArg 40
 Db 199 ACCCAACACCGAAGAAAGCTTCTCGACATCTTCACCCAGACACCGCGCAAAAGCCAGC 258
 Qy 41 ValGlnGlnIleHisGlnGlnIleuMetAlaArgLysProAlaThrLeuLysAspSerLeu 60
 Db 259 GTGGAGGAGATCCATTCAGCAGAAAGATGCTGCGAGCCCGGACCTTCMAAGACAGCTT 318
 Qy 61 GlnGlnAspLeuAsnAsnMetAsnLysPheLeuGlnLysLeuArgProLeuSerGlySer 80
 Db 319 GAGCAAGACCTCAACATATGACAAAGTCTCTGAAAAAGCTGAGGCTCTGATGGAGCC 378
 Qy 81 GluAlaProArgLeuProGlnAspProValGlyMetArgArgGlnLeuGlnLeu 100
 Db 379 GAGCTCTCTCGCGCTCCACAGACCCGCTGCGCATTCGCGCCCGCAGCTTCAGAGAGCTTG 438
 Qy 101 GlnGlnValLysAlaArgLeuGlnProTyrMetAlaGlnAlaHisGlnLeuValGlyTyr 120
 Db 439 GAGAGCTGAACCTCCCTCCACCTCCACCTCCACAGCCGCGCAGCTCTGCTGCG 498
 Qy 121 AsnLeuGlnGlyLeuArgGlnGlnIleuLysProTyrThrMetAspLeuMetGlnVal 140
 Db 499 AATTGGAGGCTTGGCGGAGCAACTGAAAGCTTCCACCAATGATGATGAGGAGAGCTG 558
 Qy 141 AlaLeuArgValGlnGlnLeuGlnGlnIleuAlaArgValValGlyLysProThrLysAla 160
 Db 559 GCCCTGCGCTCCAGAGCTCCAGAGACCTGCGCTGCGGGAAGACACCAAGCC 618
 Qy 161 GlnLeuLeuGlnGlyValLysPheGlnAlaThrAlaLeuLeuGlnGlnLeuGlnSerArgVal 180
 Db 619 CAGTTGCTGGGGGCGGGGAGAGCTTGGCTTGGAGGAGACGCGAGCCCGCTG 678
 Qy 181 ValHisHisThrGlyArgPheLysGlnLeuPheHisProTyrAlaGlnSerLeuValSer 200
 Db 679 GTGCACACACCGGCGCTCAAAAGAGCTTCCACCAATGATGATGAGGAGCTGAGTAC 738
 Qy 201 GlyIleGlyArgHisValGlnGlnLeuHisArgSerValAlaProHisAlaProAlaSer 220
 Db 739 GGCATCGGGGCGCACGTCGAGAGCTCCACCGCAGTGTGCTGCGGACCGCCCGCAGC 798
 Qy 221 ProAlaArgLeuSerArgCysValGlnValLeuSerArgLysLeuThrLeuLysAlaLys 240

```

Db      799  CCCGGCCCTAGTGGCTGCGCTGCTCTCCCGAAGCTCAGGCTCAAGGCCAAG 858
QY      241  AAlaLeuHISAlaArg11eGInGlnAsnLeuAspGlnLeuArg1uCluLeuSerArgAla 260
Db      859  GCCCTGCAGCAGCGATCCAGCAGAACTGAGACCAAGCTCGGGAGAGCTCAGCAGAGCC 918
QY      251  PheAlaGlyThrGlyThrGluGluGluValaGlyProAspProGlnMetLeuSerGluGlu 280
Db      919  TTTGACGACCTGGAGCTGACAGACGCGCGCGCCGACCCCGACAGCTCTCCGAGAG 978
QY      281  ValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyLeuGlnGlnLeuAlaPheThr 300
Db      979  GTGGCCAGCAGCTCAGGCTTCCGCCAGACCTACTGACATGACTGCTTCACT 1038
QY      301  ArgAlaIleAspGlnGluThrGluGluValaGlnGlnGlnAlaPheProProGly 320
Db      1039  CGCGCATCGACGAGAGACTGAGAGGTCCAGACAGCTGCGCCACTCCACAGAGC 1098
QY      321  HISSerAlaPheAlaProGluPheGlnGlnThrAspSerGlyLysValLeuSerLysLeu 340
Db      1099  CACACTGCTCTGCGCCAGAGTTTCAACAACAGACAGTGGCAAGCTCTGAGCAAGCTC 1158
QY      341  GlnAlaArgLeuAspAspLeuTrpGluAspIleThrHisSerLeuHisAspGlnGlyHis 360
Db      1159  CAGGCCCGCTGTGATGACTGTGGAGACATCATTCAACAGCTTCAATCACCAGGCCAC 1218
QY      361  SerHisLeuGlyAspPro 366
Db      1219  ACCCATCTGGGGGACC 1236

RESULT 7
AAF30036
ID  AAF30036 standard; DNA; 1879 bp.
AC
XX
AC  AAF30036;
XX
DT  23-APR-2001 (first entry)
XX
DE  Human apolipoprotein A-IV-related protein (AA4RP) cDNA.
XX
KW  Apolipoprotein A-IV-related protein; AA4RP; human;
    biallelic marker; lipid metabolism; liver related disorder;
    diagnosis; gene therapy; obesity; diabetes; coronary heart disease;
    ss.
XX
OS  Homo sapiens.
XX
FH  Key Location/Qualifiers
FT  1..20
FT  5'UTR /*tag= a
FT  21..1121
FT  CDS /*tag= b
FT  3'UTR 1122..1879
FT  /*tag= c
FT  allele replace(1153,C,T)
FT  /*tag= d
FT  /note= "biallelic marker 17-42-250, polymorphic
        base C or T"
XX
XX  MO200100803-A2.
XX
XX  04-JAN-2001.
XX
XX  21-JUN-2000; 2000MO-IB01011.
XX
XX  25-JUN-1999; 99US-0141032.
XX  20-DEC-1999; 99MO-IB02058.
XX  21-DEC-1999; 99US-0469099.
XX
XX  (GEST ) GENSET.
XX
PI  Yen F, Denison B, Bour B, Bihain B, Bougueleret L, Duclet A;

```

```

PI  Dumas Mline Edwards J;
XX
XX  WPI: 2001-071485/08.
DR  P-PSDB; AAB20103.
XX
PT  Nucleic acids encoding apolipoprotein A-IV-related proteins (AA4RP) and
PT  biallelic markers of AA4RP, useful for diagnosing lipid metabolism
PT  related disorders and/or liver related disorders.
PS  Claim 5; Page 238-239; 260pp; English.
XX
XX  The present sequence is that of cDNA encoding human apolipoprotein
CC  A-IV-related protein (AA4RP, see AAB20103). AA4RP cDNA clone
CC  117-005-2-0-E10-FLC is deposited as ECACC 99061735. AA4RP is
CC  differentially expressed in obese mouse models, indicating a role
CC  in lipid metabolism disorders. It appears to be the human homologue
CC  of rat regeneration associated protein (RAP3), which is believed to
CC  be involved in liver regeneration, and is likely to have a similar
CC  function. It also shows 52% similarity to apolipoprotein A-IV, and
CC  is likely to have a similar function. The invention also provides
CC  the AA4RP gene, and methods for genotyping a nucleic acid
CC  containing 1 or more of the biallelic markers. Also provided is a
CC  method for detecting a statistical correlation between a biallelic
CC  marker allele and a phenotype. Diagnostic methods are provided for
CC  determining whether an individual is at risk of developing a lipid
CC  metabolism related disorder and/or a liver related disorder, such
CC  as obesity, diabetes and coronary heart disease, and whether the
CC  individual suffers from such a disorder as a result of a
CC  polymorphism in the AA4RP gene. AA4RP DNA constructs can be used
CC  to direct temporal and spatial AA4RP gene expression in recombinant
CC  cell hosts and in transgenic animals. They can also be used for
CC  recombinant production of AA4RP polypeptides, and as probes and
CC  primers.
XX
SQ  Sequence 1879 BP; 395 A; 540 C; 590 G; 353 T; 1 other:

Alignment Scores:
Pred. No.: 0 Length: 1879
Score: 366.00 Matches: 366
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-842-364-3 (1-366) x AAF30036 (1-1879)
QY 1 MetAlaSerMetAlaValLeuThrTrpAlaLeuLeuSerAlaPheSerAla 20
Db 21 ATGCGAAGCATGGCTGCCCTGCCCTGCTCTCTTTTCACCGCTTTGCGCC 80
QY 21 ThrGlnAlaArgLysGlyPheTrpAspTyrPheSerGlnThrSerGlyAspLysGlyArg 40
Db 81 ACCCAGGACGAGAAAGGCTTCTGTGGACTACTTACGACAGCAGCGGGGACAAAGCCAGG 140
QY 41 ValGluGlnIleHisGlnGlnLysMetAlaArgGluProAlaThrLeuLysAspSerLeu 60
Db 141 GTGCAGCAGATCATCAGCAGAGAGATGCTCGGACCCGCCGCTGAAGACACACCTTT 200
QY 61 GluGlnAspLeuAsnAsnMetLysPheLeuGluLysLeuArgProLeuSerGlySer 80
Db 201 GAGCAAGACCTCAACAATATGAAAGATTCTCTGGAAAAAGCTCAGGCTGTGACTGGAGCC 260
QY 81 GluAlaProArgLeuProGlnAspProValaGlyMetArgArgGlnLeuGlnGluLeu 100
Db 261 GAGGCTCTCGGCTCCACAGAGACCGGTGGGATCGCGCGGAGCTGAGAGGAGTTG 320
QY 101 GluGluValLysAlaArgLeuGlnProTyrMetAlaGlnAlaHisGluLeuValGlyTyr 120
Db 321 GAGGAGGTGAAGGCTGCTCCAGCCCTACATGCGACAGCGCCACAGCTGTGGGCTGG 380
QY 121 AsnLeuGluGlnLeuArgGlnGlnLeuLysProTyrThrMetAspLeuMetGluGlnVal 140

```



```

Db 321 GAGGAGTGAAGGCTCGCTCCAGCCCTACATGGACAGGCGACGAGCTGGTGGCTGG 380
QY 121 AsnLeuGluGlyLeuArgGlnGlnLeuLysProTyrThrMetAspLeuMetGlnGlnVal 140
Db 381 AATTGGAGGGCTTGGGAGCAGCAACTGAAGCCCTACACGATGATGTGATGGAGCAGGTG 440
QY 141 AlaLeuArgValGlnGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 441 GCCCTCGCCGCTGACAGAGCTGACAGACACTTCCGCTGGGGAAGACACCAAGCC 500
QY 161 GlnLeuLeuGlyGlyValAspGlnAlaTPrAlaLeuLeuGlnGlnGlnGlnGlnGlnGln 180
Db 501 CAGTTCCTGGGGGCGGTGGACGAGGCTTGGCTTGTCTGAGGAGTCTGACGAGCCGCTG 560
QY 181 ValHisHisThrGlyArgPheLysGlyLeuPheHisProTyrAlaGlySerLeuValSer 200
Db 561 GTGCACCAACACGGCGCTTCACAAAGAGCTTCCACCCATACCCGAGAGCTGGTAGC 620
QY 201 GlyTLeGlyArgHisValGlnGlnLeuHisArgSerValAlaProHisAlaProAlaSer 220
Db 621 GCGATCGGGCGCCACCTGCACGACGACTCCACCGAGTGTGCTCCGACGCGCCCGCAGC 680
QY 221 ProAlaArgLeuSerArgGlyValGlnValLeuSerArgLysLeuThrLeuLysAlaLys 240
Db 681 CCCGCGCGCTCAGTGTGCTGCTGAGTGTCTCCCGAAGCTCACGCTCAAGGCCAAG 740
QY 241 AlaLeuHisAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
Db 741 GCGCTCCACGACGACATCCACGACGACGACGACGACGACGACGACGACGACGACGACG 800
QY 261 PheAlaGlyThrGlyThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 801 TTTTGCAGGACACTGGGACTGACGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860
QY 281 ValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeuGlnGlnGlnGlnGlnGln 300
Db 861 GTTCCGCGACGACGACTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 920
QY 301 ArgAlaIleAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
Db 921 CCGCGCATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 980
QY 321 HisSerAlaPheAlaProGlnPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 340
Db 981 CACAGTCCCTGCGCCGACAGTTTCAACAACACACAGAGTCTGACAGCAAGCTG 1040
QY 341 GlnAlaArgLeuAspLeuThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db 1041 CAGGCCCGCTGTGATGACCTGTGGGAGACATCTACACAGCTTCTATGACGAGGCCAC 1100
QY 361 SerHisLeuGlyAspPro 366
Db 1101 AGCCATCTGGGGGACCCC 1118

```

RESULT 9
AAF64022
ID AAF64022 standard; cDNA; 1894 bp.

AC AAF64022;
DT 05-APR-2001 (first entry)

XX cDNA encoding human secreted protein #23.

KW Secreted protein; prevention; treatment; diagnosis; disease;

OS infection; ds.

XX Homo sapiens.

XX WO200100806-A2.

PN 04-JAN-2001.

```

XX 21-JUN-2000; 2000MO-IB00951.
PF 25-JUN-1999; 99US-0141032.
PR 21-DEC-1999; 99US-0469099.
XX (GEST) GENSET.
PA Dumas Milne Edwards J, Houguetier L, Jobert S;
PI MPI; 2001-071487/08.
DR 49 Secreted proteins and the nucleic acids encoding them, useful in
PT gene therapy and for detecting similar sequences in samples.
PS Claim 1; Page 234-235; 307pp; English.
XX The present invention relates to 49 secreted proteins and the cDNAs
CC encoding them. The protein and nucleic acids may be used in the
CC prevention, treatment and diagnosis of diseases associated with
XX inappropriate protein expression.
SQ Sequence 1894 bp; 411 A; 541 C; 588 G; 353 T; 1 other:
Alignment Scores:
Pred. No.: 0 Length: 1894
Score: 366.00 Matches: 366
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0
US-09-842-364-3 (1-366) x AAF64022 (1-1894)
QY 1 MetAlaSerMetAlaIaValIleuThrTrpAlaLeuAlaLeuLeuSerAlaPheSerAla 20
Db 21 ATGCAAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 80
QY 21 ThrGlnAlaArgLysGlyPheThrPaspTyrPheSerGlnThrSerGlnLysGlnGlnGln 40
Db 81 ACCCAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 140
QY 41 ValGlnGlnIleHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
Db 141 GTGACACCATTCATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 200
QY 61 GlnGlnAspLeuAsnAsnMetAlaLysPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
Db 201 GAGCAAGACCTTCAACATATGAAAGTTCTGGAAGCTGAGGCTTGAAGTGGAGAC 260
QY 81 GluAlaProArgLeuProGlnAspProValGlyMetArgArgGlnGlnGlnGlnGlnGln 100
Db 261 GACCTCTCTCCGCTCCACAGCAGCCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 320
QY 101 GlnGlnValLysAlaArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
Db 321 CAGGAGCTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
QY 121 AsnLeuGluGlyLeuArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
Db 381 AATTGGAGGGCTTGGGAGCAGCAACTGAAGCCCTACACGATGATGTGATGGAGCAGGTG 440
QY 141 AlaLeuArgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 441 GCCCTCGCCGCTGACAGAGCTGACAGACACTTCCGCTGGGGAAGACACCAAGCC 500
QY 161 GlnLeuLeuGlyGlyValAspGlnAlaTPrAlaLeuLeuGlnGlnGlnGlnGlnGlnGln 180
Db 501 CAGTTCCTGGGGGCGGTGGACGAGGCTTGGCTTGTCTGAGGAGTCTGACGAGCCGCTG 560
QY 181 ValHisHisThrGlyArgPheLysGlyLeuPheHisProTyrAlaGlySerLeuValSer 200
Db 561 GTGCACCAACACGGCGCTTCACAAAGAGCTTCCACCCATACCCGAGAGCTGGTAGC 620

```

QY 201 GYILEGLYARHISValGIngluLeuHISARSerValAlaProHISAlaProAlaSer 220
DB 621 GGCATGGGGCCACGCGAGAGCTGCACCGAGTGTGCTCCGACGCGCCGCGACG 680
QY 221 PROAlaArgLeuSerArgCysValGInValLeuSerArgCysLeuThrLeuLysAlaLys 240
DB 681 CCGGCGGCGCTGATGGCTGCTGCGAGTGTCTCTCCGGAAGTCAACGCTCAAGGCCAAG 740
QY 241 AlAlaLeuHISAlaArgGIngluLeuHISARSerValAlaProHISAlaProAlaSer 260
DB 741 GCCCTCAGCGACGCACTCCAGCAAGCTGAGACACTGCGCAAGAGCTCAGAGAGCC 800
QY 261 PheAlaGlyThrGlyThrGlyGluGlyAlaGlyProAspProGlnMetLeuSerGluGlu 280
DB 801 TTGTGAGGCGACTGCGACTGAGCAAGGCGCGCCGACCCGACAGATGCTCTCCGAGAG 860
QY 281 ValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeuGlnIleAlaAlaPheThr 300
DB 861 GTGCGCGAGGACTTCAAGGCTTCCGCGACAGACCTACTGAGATAGTGCCTTCACT 920
QY 301 ArgAlaIleAspGlnGluThrGluGluValGInGlnGlnLeuAlaProProProGly 320
DB 921 CGGCGCATGCGACGAGAGACTGAGAGGTCCACAGCAGCTGCGCACCTCCACAGCG 980
QY 321 HISSerAlaPheAlaProGluPheGlnGlnThrAspSerGlyLysValLeuSerLysLeu 340
DB 981 CACAGTGGCTTCCGCGACAGTTTCAACAACAGACAGTGGCAAGGTTCTGACCAAGCTG 1040
QY 341 GlnAlaArgLeuAspAspLeuTrpGluAspIleThrHISerLeuHISAspGlnGlyHis 360
DB 1041 CAGGCGCGTGTGATGACCTGTGGGAAGACATCATCAGCTTATGACCAAGGCGCAC 1100
QY 361 SerHISLeuGlyAspPro 366
DB 1101 AGCCATCTGGGGAGCCCC 1118

RESULT 10
AAK89395
ID AAK89395 standard: DNA; 2249 bp.
XX
AC AAK89395;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 2971.
XX
KM Human; digestive system antigen; gene therapy; cancer; appendicitis;
KM ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KM digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214866.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI: 2001-502630/55.
 XX
 XX Polynucleotides encoding digestive system antigens, useful for
 XX diagnosing, treating, preventing and/or prognosing disorders of the
 XX digestive system, particularly cancer and cancer metastases -
 XX
 XX Disclosure; SEQ ID NO 2971: 986pp; English.
 XX
 XX The present invention provides the protein and coding sequences of a
 XX number of human digestive system antigens. These can be used in the
 XX diagnosis, treatment and prevention of digestive system disorders,
 XX including cancer, Meckel's diverticulum, bacterial or parasitic
 XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
 XX ulcerative colitis. The present sequence is a genomic DNA fragment
 XX encoding a digestive system antigen of the invention.
 XX
 XX Sequence 2249 BP; 427 A; 671 C; 717 G; 434 T; 0 other;
 XX

Pred. No.: 3,026-294 Length: 2249
 Score: 312.00 Matches: 312
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.25% Indels: 0
 DB: 22 Gaps: 0
 US-09-842-364-3 (1-366) x AA089395 (1-2249)
 QY 55 ThrLeuLysAspSerLeuGluGlnAspLeuAsnMetAsnLysPheLeuGluLysIleu 74
 DB 857 ACCCTGAAGACAGCCTTGAAGACCTCAACATATGACAGATCTCGGAAAGAGCTG 916
 QY 75 ArgProLeuSerGlySerGluAlaProArgLeuProGlnAspProValGlyMetArgArg 94
 DB 917 AGCCCTCTGACTGCGAGCAGGCTCCCTCCGCTCCACAGACCCCGCTGGACATCCGCGCG 976
 QY 95 GlnLeuGlnGluGluLeuGluGluValLysAlaArgLeuGlnProArgMetAlaGluAla 114
 DB 977 CAGCTCAGAGAGAGATTGGAGAGGTGAAGGCTGCGCTCCAGGCTTACATGAGCAGAGGGC 1036
 QY 115 HisGluLeuValGlyTrpAsnLeuGluGlyIleuArgGlnGlnLeuLysProArgTrpMet 134
 DB 1037 CAGGAGCTGATGGGCTGGGAATTTGGAGGGCTTGGGAGCACTGAAGCCCTACAGAGT 1096
 QY 135 AspLeuMetGluGlnValAlaLeuArgValGlnGluLeuGlnGluGlnLeuArgVal 154
 DB 1097 GATCTCATGAGAGAGCTGCGCTCCGCTCCAGCAGCTCAGAGCAGATTCGCGCTGTG 1156
 QY 155 GlyGluAspThrLysAlaGlnLeuLeuGlyValAspGluAlaTrpAlaLeuGln 174
 DB 1157 GGGGAAGACACCAAGGCCCAAGTTGCTGGGGGGGTGGAGAGGCTTGGCTTGTGTCAG 1216
 QY 175 GlyLeuGlnSerArgValAlaHisIsthrcIyArgPheLysGluLeuPheHisProTyr 194
 DB 1217 GGACTGCAGAGCGCGCTGTGTGCACACCGCGCGCTTCAAGAGCTCTCCACCCATAC 1276
 QY 195 AlaGluSerLeuValSerGlyIleGlyAlaGlnIsthrcIyLeuHisArgSerValAla 214
 DB 1277 GCCAGAGCCTGCTGAGCGGCATCCGCGCCACGCTCAGAGCTCAGCCGCGCTGTGCT 1336
 QY 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
 DB 1337 CGGACGCGCCCGCCAGCGCCCGCCGCTCAGTGTGTGCTGCTGCTGCTGCTGCTGCTG 1396
 QY 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnAsnLeuAspGlnLeuArg 254
 DB 1397 CTCACGCTCAAGGCCAAGGCTCTGCACGACGATCCAGACCAACCTGAGCAGCTGGCC 1456
 QY 255 GluGluLeuSerArgAlaPheAlaGlyThrGlyThrGluGluValAlaGlyProAspPro 274
 DB 1457 GAAGAGCTCAGCAGAGCTTTCAGGCACTGCGACTGAGAGAGGGCGCCGCGGAGCCCC 1516
 QY 275 GlnMetLeuSerGluGluValArgGlnArgLeuGlnAlaPheArgGlnAspThrArgLeu 294
 DB 1517 CAGATGCTCTCCGAGAGAGGTGGCCAGGACACTTTCAGGCTTTCGCGCAGAGACCTACTG 1576
 QY 295 GlnIleAlaIaPheThrArgAlaIleAspGlnGluThrGluGluValGlnGlnLeu 314
 DB 1577 CAGATAGCTGCTTCACTGCGCCATGACACCGAGAGCTGAGAGAGTCCAGAGCAGCTG 1636
 QY 315 AlaProProProGluHisSerAlaPheAlaProGluPheGlnGlnThrAspSerGly 334
 DB 1637 GCGCCACTCCACAGGCGACAGTCTTCCGCCAGAGTTTCAACAAACAGACAGTGGC 1696
 QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGlnAspIleThrHisSer 354
 DB 1697 AAGGTTCAGACAAGCTGAGGCGGCTGTGATGACCTGTGGGAAGACATCACTACAGC 1756
 QY 355 LeuHisAspGlnGlyHisSerHisLeuGlyAspPro 366
 DB 1757 CTTCAATACACAGGCGCACAGCATCTGTGGGAGACCC 1792

Alignment Scores:

```
RESULT 11
AAK89396
ID AAK89396 standard; DNA; 2250 BP.
XX
AC AAK89396;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 2972.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
```

```
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244611.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
```

PR 01-DEC-2000: 2000US-0250391.
 PR 05-DEC-2000: 2000US-0251030.
 PR 05-DEC-2000: 2000US-0251988.
 PR 05-DEC-2000: 2000US-0256719.
 PR 06-DEC-2000: 2000US-0251479.
 PR 08-DEC-2000: 2000US-0251856.
 PR 08-DEC-2000: 2000US-0251868.
 PR 08-DEC-2000: 2000US-0251869.
 PR 08-DEC-2000: 2000US-0251989.
 PR 08-DEC-2000: 2000US-0251990.
 PR 11-DEC-2000: 2000US-0254097.
 PR 05-JAN-2001: 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Roscon CA, Barash SC, Ruben SM;

XX WPI: 2001-502630/55.

PT Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognosing disorders of the
 PT digestive system, particularly cancer and cancer metastases -

XX Disclosure: SEQ ID NO 2972; 986bp; English.

XX The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a genomic DNA fragment
 CC encoding a digestive system antigen of the invention.

XX Sequence 2250 BP: 427 A; 674 C; 716 G; 433 T; 0 other:

Alignment Scores:
 Pred. NO.: 3.02e-294 Length: 2250
 Score: 312.00 Matches: 312
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 85.25% Indels: 0
 DB: 22 Gaps: 0

US-09-842-364-3 (1-366) x AAK89396 (1-2250)

QY 55 ThrLeuLysASPserLeuGlnGlnAspLeuAsnMetAsnLysPheLeuGlnLysLeu 74
 DB 858 ACCCTGAAGAGAGCTTGAGCAGACCTCAACATATGACAAAGTCTCTGAAAAGCTG 917
 QY 75 ArgProLeuSerGlySerGluAlaProArgLeuProGlnAspProValGlyMetLArgArg 94
 DB 918 AGGCTCTGAGTGGGAGCGAGCGCTCTCGCCCTCCACAGACCGCTGGCGCATGCCCGG 977
 QY 95 GlnLeuGlnGlnGlnLeuGlnGlnValLysAlaArgLeuGlnProTyrMetAlaGlnAla 114
 DB 978 CAGCTGAGGAGAGAGTGGAGAGGTGAAGGCTCGCTCCAGCCCTACATGACAGAGCG 1037
 QY 115 HisGluLeuValGlyTyrAsnLeuGlnGlnLysLeuArgGlnGlnLeuLysProTyrThrMet 134
 DB 1038 CACGAGCTGGTGGCTGGAATTTGGAGGGCTTGGCGCAGCACTAAGCCCTACACGATG 1097
 QY 135 AspleuMetGluGlnValAlaLeuArgValGlnGlnLeuGlnGlnLeuArgValAla 154
 DB 1098 GATCTGATGAGAGAGCTGGCCCTGGCGCTGCAGAGCTGCAGACAGCTGGCGCTGGCG 1157
 QY 155 GlyGluAspTrpTrpLysAlaGlnLeuLeuGlnGlnValLysAlaArgValAlaTrpAlaLeuGln 174
 DB 1158 GGGGAAAGACCAAGGCCAGCTGCTGGGGGGCGGAGGAGGCTTGGGCTTGTCTGCAG 1217
 QY 175 GlyLeuGlnSerArgValAlaHisIshTrpArgPheLysGlnLeuPheHisProTyr 194
 DB 1218 GGACTGCGAGAGCCGGCTGTGTGCACACACCGCCCTTCAAGAGCTCTTCCACCCATAC 1277

QY 195 AlaGluSerLeuValSerGlyIleGlyArgHisValGlnGlnLeuHisArgSerValAla 214
 DB 1278 GCCGAGAGCTGCTGAGCGGCATCGGGCCCGCAGCTGCAGAGCTGCACCCAGCTGGCT 1337
 QY 215 ProHisAlaProHisArgProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
 DB 1338 CCGCAGCCCCCGCCAGCCCCGGCGGCTGAGTGGCTGGAGGAGGCTCTCCCGGAG 1397
 QY 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnAsnLeuAspGlnLeuArg 254
 DB 1398 CTCAGCCTCAAGGCCCAAGGCCCTTCAGCCAGCCAGCATCAGAGCAAGCTGCACCTGCC 1457
 QY 255 GlnGluLeuSerArgAlaPheAlaGlyThrGlyThrGlnGlnGlnValAlaGlyProAspPro 274
 DB 1458 GAAGACCTCAGCAGAGCTTTCGAGGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1517
 QY 275 GlnMetLeuSerGlnGlnValArgGlnArgLysGlnAlaPheArgGlnAspTrpTyrLeu 294
 DB 1518 CAGATCTCTCCAGAGAGGTGGCGCAGCAGCTTCAAGCTTTCGCGCAGACACTACTG 1577
 QY 295 GlnIleAlaAlaPheThrArgAlaIleAspGlnGlnGlnGlnValGlnGlnGlnLeu 314
 DB 1578 CACATAGCTGCTTTCACCTCGCGCCATCGACACAGACACTGAGAGAGCTCCAGCAGCAGCT 1637
 QY 315 AlaProProProProGlnHisSerAlaPheAlaProGlnGlnGlnGlnHisArgSerGly 334
 DB 1638 GCCCCACCTTCACACAGCCACACACTGCTTCCGCCAGCTTTCACCAACACAGCAGCTGCC 1697
 QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGlnAspIleThrHisSer 354
 DB 1698 AAGGTTCTAGCAGAGTGCAGAGGCCCTGTGATGACCTGTGGAGAACATCACTACAGC 1757
 QY 355 LeuHisAspGlnGlnLysHisSerHisLeuGlnLysAspPro 366
 DB 1758 CTTCATGACACAGAGGCCACAGCCATCTGGGGAGCC 1793

RESULT 12
 ID AAK89397 standard; DNA; 2250 BP.
 XX AAK89397;
 AC
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human digestive system antigen genomic sequence SEQ ID NO: 2973.
 XX
 KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 KW digestive system disorder; Meckel's diverticulum; ds.
 OS Homo sapiens.
 PN WO200155314-A2.
 PD
 XX
 XX 02-AUG-2001.
 PF 17-JAN-2001; 2001WO-US01324.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0188874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0232968.
 PR 14-SEP-2000; 2000US-0232969.
 PR 14-SEP-2000; 2000US-0232997.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0233400.
 PR 14-SEP-2000; 2000US-0233401.
 PR 14-SEP-2000; 2000US-0233663.
 PR 14-SEP-2000; 2000US-0233664.
 PR 21-SEP-2000; 2000US-0233665.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234984.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 20-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DIC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2000US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Harash SC, Ruben SM;
 PI WPI: 2001-502630/55.
 XX
 PT Polynucleotides encoding digestive system antigens, useful for
 PT diagnosing, treating, preventing and/or prognostizing disorders of the
 PT digestive system, particularly cancer and cancer metastases -
 XX
 PS Disclosure: SEQ ID NO 2973; 986pp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosing, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a genomic DNA fragment
 CC encoding a digestive system antigen of the invention.
 XX
 SQ Sequence 2250 BP; 426 A; 673 C; 718 G; 433 T; 0 other;
 Alignment Scores: 3.02e-294 Length: 2250
 Pred. No.: 2250

FT	/note= "17-40.rp"
FT	2964..13984
FT	/*tag= t
FT	/note= "17-41.pu"
FT	complement (3432..14454)
FT	/*tag= u
FT	/note= "17-41.rp"
FT	300..318
FT	/*tag= v
FT	/note= "17-42-319.mis"
FT	320..338 (42572..42591)
FT	/*tag= w
FT	/note= "17-42-319.mis"
FT	3194..3212
FT	/*tag= z
FT	/note= "17-41-250.mis"
FT	complement (3214..3232)
FT	/*tag= y
FT	/note= "17-41-250.mis"
FT	307..331
FT	/*tag= z
FT	/note= "17-42-319 probe
FT	3201..3225
FT	/*tag= aa
FT	/note= "17-41-250 probe"
FT	

PN	WO200100803-A2.
XX	
PD	04-JAN-2001.
XX	
PF	21-JUN-2000; 2000WO-1B01011.
XX	
PR	25-JUN-1999; 99US-0141032.
PR	20-DEC-1999; 99WO-1B02058.
PR	21-DEC-1999; 99US-0469099.
XX	
PA	(GRST) GENSET.
XX	
PI	Yen F, Denison B, Bour B, Blhain B, Bouguetelarel L, Duclert A;
PI	Dunnas Milne Edwards J;
XX	
DR	WPI; 2001-071485/08.

used to direct temporal and spatial AARF gene expression in recombinant cell hosts and in transgenic animals. They can also be used for recombinant production of AARF polypeptides, and as

OY		355	LeuHISASPGINGLYHISserNILEUGLYAspro	366				
D8		3144	CITTCATGACAGCGCCACAGCACACTGGGGACC	3179				
RESULT 14								
ID	AAF30035	AAAF30035 standard; DNA: 81001 BP.						
XX	AAF30035;							
DT	23-APR-2001	(first entry)						
DE	Human apolipoprotein A-IV-related protein (AARF) gene.							
KW	Apolipoprotein A-IV-related protein; AARF; human;							
RW	bi allelic marker; lipid metabolism; liver related disorder;							
KM	obesity; diabetes; coronary heart disease; diagnosis; gene therapy;							
XX	chromosome 11: ds.							
OS	Homo sapiens.							
Key	Location/Qualifiers							
FT	5'UTR	10946..12946						
FT		/tag= a						
FT	CDS	13478..15210						
FT		/tag= b						
FT	exon	12947..12958						
FT		/tag= c						
FT	intron	/number= 1 12959..13469						
FT		/tag= d						
FT	exon	/number= 1 13470..13526						
FT		/tag= e						
FT	intron	/number= 2 13527..13640						
FT		/tag= f						
FT	exon	/number= 2 13641..13752						
FT		/tag= g						
FT	intron	/number= 3 13753..14270						
FT		/tag= h						
FT	exon	/number= 3 14271..15968						
FT		/tag= i						
FT	3'UTR	/number= 4 15969..17969						
FT		/tag= j						
FT	allele	replace(1239,C,T)						
FT		/tag= k						
FT	allele	/note= "biallelic marker 20-828-311, polymorphic base C or T"						
FT	allele	replace(12347,C,T)						
FT		/tag= l						
FT	allele	/note= "biallelic marker 17-42-319, polymorphic base C or T"						
FT	allele	replace(15241,C,T)						
FT		/tag= m						
FT	allele	/note= "biallelic marker 17-41-250, polymorphic base C or T"						
FT	allele	replace(42218,A,G)						
FT		/tag= n						
FT	allele	/note= "biallelic marker 20-841-149, polymorphic base A or G"						
FT	allele	replace(45442,A,G)						
FT		/tag= o						
FT	allele	/note= "biallelic marker 20-842-115, polymorphic base A or G"						
FT	allele	replace(77058,C,T)						
FT		/tag= p						
FT	allele	/note= "biallelic marker 20-853-415, polymorphic base A or G"						

FT	primer_bind	929..949	base C or T"
FT		/*tag= q	
FT		/note= "primer 20-828.pu"	
FT	primer_bind	complement (1357..1377)	
FT		/*tag= r	
FT		/note= "primer 20-828.rp"	
FT	primer_bind	12059..12050	
FT		/*tag= s	
FT		/note= "primer 17-42.pu"	
FT	primer_bind	complement (12581..12603)	
FT		/*tag= t	
FT		/note= "primer 17-42.rp"	
FT	primer_bind	14992..15012	
FT		/*tag= u	
FT		/note= "primer 17-41.pu"	
FT	primer_bind	complement (15460..15482)	
FT		/*tag= v	
FT		/note= "primer 17-41.rp"	
FT	primer_bind	42070..42090	
FT		/*tag= w	
FT		/note= "primer 20-841.pu"	
FT	primer_bind	complement (42572..42591)	
FT		/*tag= x	
FT		/note= "primer 20-841.rp"	
FT	primer_bind	45328..45347	
FT		/*tag= y	
FT		/note= "primer 20-842.pu"	
FT	primer_bind	complement (45863..45883)	
FT		/*tag= z	
FT		/note= "primer 20-842.rp"	
FT	primer_bind	76644..76664	
FT		/*tag= aa	
FT		/note= "primer 20-853.pu"	
FT	primer_bind	complement (77166..77185)	
FT		/*tag= ab	
FT		/note= "primer 20-842.rp"	
FT	primer_bind	12220..1238	
FT		/*tag= ac	
FT		/note= "primer 20-828-311.mis"	
FT	primer_bind	complement (1240..1258)	
FT		/*tag= ad	
FT		/note= "primer 20-828-311.mis"	
FT	primer_bind	12328..12346	
FT		/*tag= ae	
FT		/note= "primer 17-42-319.mis"	
FT	primer_bind	complement (12348..12366)	
FT		/*tag= af	
FT		/note= "primer 17-42-319.mis"	
FT	primer_bind	15222..15240	
FT		/*tag= ag	
FT		/note= "primer 17-41-250.mis"	
FT	primer_bind	complement (15242..15260)	
FT		/*tag= ah	
FT		/note= "primer 17-41-250.mis"	
FT	primer_bind	42199..42237	
FT		/*tag= ai	
FT		/note= "primer 20-841-149.mis"	
FT	primer_bind	complement (42219..42237)	
FT		/*tag= aj	
FT		/note= "primer 20-841-149.mis"	
FT	primer_bind	45423..45441	
FT		/*tag= ak	
FT		/note= "primer 20-842-115.mis"	
FT	primer_bind	complement (45443..45461)	
FT		/*tag= al	
FT		/note= "primer 20-842-115.mis"	
FT	primer_bind	77039..77057	
FT		/*tag= am	
FT		/note= "primer 20-853-415.mis"	
FT	primer_bind	complement (77059..77077)	
FT		/*tag= an	
FT		/note= "primer 20-853-415.mis"	

Fri Jan 24 10:09:13 2003

us-09-842-364-3401ig.rmg

Page 21

Search completed: January 23, 2003, 18:48:57
Job time : 289 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 23, 2003, 17:47:27 (Search time 54 Seconds
(Without alignments)
2078.587 Million cell updates/sec)

Title: US-09-842-364-3
Sequence: 1 MASMAVLTWALALSAFSA.....LMEDITHSLDQSHLGD 366

Scoring table:
OLIGO
Xgapop 60.0, Ygapext 60.0
Ygapop 60.0, Ygapext 60.0
Ygapop 6.0, Ygapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 segs, 153338381 residues

Word size: 6

Total number of hits satisfying chosen parameters: 4253

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.p2n.model -DEV=x1h
-O/cgn2.1/USPTO.spool/US09842364/runat_16012003_164829_10671/app_query.fasta.1.519
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=olig.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITs=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cd1
-LIST=45 -DOCALLIN=200 -THR SCORE=quality -THR MIN=6 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09842364 @cgn2.1.1.25 @runat_16012003_164829_10671 -NCPu=6 -ICPu=3
-NO_XIPXY -NO_MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -IONCLOG -DEV=TIMEOUT=120
-WARN=TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FEAPOP=6 -FEAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELop=6 -DELEXT=7

Database:

Issued_Patents_NA.*
1: /cgn2.6/prodata/2/ina/5A.COMB.seq.*
2: /cgn2.6/prodata/2/ina/5A.COMB.seq.*
3: /cgn2.6/prodata/2/ina/6A.COMB.seq.*
4: /cgn2.6/prodata/2/ina/6A.COMB.seq.*
5: /cgn2.6/prodata/2/ina/6A.COMB.seq.*
6: /cgn2.6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	366	100.0	1879	4	US-09-750-580-2
2	312	85.2	5361	4	US-09-750-580-4
3	312	85.2	81001	4	US-09-750-580-1
4	176	48.1	592	4	US-09-276-531-115
5	48	2.2	48	4	US-08-487-811A-23
6	8	2.2	48	4	US-09-060-694-23
7	8	2.2	105	2	US-08-623-471-6
8	8	2.2	201	2	US-08-623-471-7
9	8	2.2	345	2	US-08-623-471-8
10	8	2.2	397	1	US-08-592-126-106
11	8	2.2	409	1	US-07-626-618A-8
12	8	2.2	409	1	US-08-333-977-8

13	8	2.2	563	1	US-07-928-611-8	Sequence 8, Appl1
14	8	2.2	563	4	US-08-487-811A-8	Sequence 8, Appl1
15	8	2.2	563	4	US-09-060-694-8	Sequence 8, Appl1
16	8	2.2	563	5	US-09-378-074-8	Sequence 8, Appl1
17	8	2.2	659	2	PCT-US93-07370-8	Sequence 10, Appl1
18	8	2.2	659	2	US-07-928-611-10	Sequence 10, Appl1
19	8	2.2	659	2	US-08-487-811A-10	Sequence 10, Appl1
20	8	2.2	659	4	US-09-060-694-10	Sequence 10, Appl1
21	8	2.2	659	4	US-09-378-074-10	Sequence 10, Appl1
22	8	2.2	659	5	PCT-US93-07370-10	Sequence 10, Appl1
23	8	2.2	687	4	US-09-040-984-38	Sequence 38, Appl1
24	8	2.2	687	4	US-09-123-912-38	Sequence 38, Appl1
25	8	2.2	687	4	US-09-643-597-38	Sequence 38, Appl1
26	8	2.2	745	1	US-08-036-5558-163	Sequence 163, App
27	8	2.2	745	1	US-08-469-569-163	Sequence 163, App
28	8	2.2	745	1	US-08-249-322A-163	Sequence 163, App
29	8	2.2	745	1	US-08-469-526A-163	Sequence 163, App
30	8	2.2	745	2	US-08-734-591A-163	Sequence 163, App
31	8	2.2	745	2	US-08-469-660-163	Sequence 163, App
32	8	2.2	745	3	US-08-341-018-51	Sequence 51, Appl
33	8	2.2	745	3	US-08-470-335-163	Sequence 163, App
34	8	2.2	745	4	US-08-735-021-163	Sequence 163, App
35	8	2.2	745	4	US-08-734-664A-163	Sequence 163, App
36	8	2.2	745	4	US-08-470-339-163	Sequence 163, App
37	8	2.2	745	4	US-08-467-602-163	Sequence 163, App
38	8	2.2	745	5	PCT-US94-05083C-159	Sequence 159, App
39	8	2.2	745	5	PCT-US95-06846A-163	Sequence 163, App
40	8	2.2	763	4	US-09-615-192A-117	Sequence 117, App
41	8	2.2	777	2	US-08-975-316-37	Sequence 37, Appl
42	8	2.2	803	1	US-09-615-192A-37	Sequence 37, Appl
43	8	2.2	803	1	US-07-928-611-12	Sequence 12, Appl
44	8	2.2	803	2	US-08-487-811A-12	Sequence 12, Appl
45	8	2.2	803	4	US-09-060-694-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-750-580-2
Sequence 2, Application US/09750580
Patent No. 6453280
GENERAL INFORMATION:
APPLICANT: Yen, Frances
APPLICANT: Denison, Blake
APPLICANT: Bour, Barbara
APPLICANT: Bitrain, Bernard
APPLICANT: Dumas Mline Edwards, Jean-Baptiste
APPLICANT: Duclert, Aymeric
APPLICANT: Bouqueleret, Lydie
APPLICANT: Ebbels-Reed, Dana
APPLICANT: Saitor-Cid, Luisa
TITLE OR INVENTION: METHODS AND COMPOSITIONS FOR INHIBITTING NEOPLASTIC CELL GROW
FILE REFERENCE: 89.052.CIP
CURRENT APPLICATION NUMBER: US/09750580
CURRENT FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: US 09/599,362
PRIOR FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: PCT/IB00/0101
PRIOR FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: PCT/IB99/02058
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: US 49/469/099
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: US 60/113,686
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: US 60/141,032
PRIOR FILING DATE: 1999-06-25
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent.pm
SEQ ID NO 2
LENGTH: 1879
TYPE: DNA
ORGANISM: Homo sapiens

```

FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..29
: NAME/KEY: CDS
: LOCATION: 30..1121
: NAME/KEY: 3'UTR
: LOCATION: 1122..1879
: NAME/KEY: allele
: LOCATION: 1153
: OTHER INFORMATION: 17-41-250
US-09-750-580-2 : polymorphic base C or T

Alignment Scores:
Pred. No.: 0 Length: 1879
Score: 366.00 Matches: 366
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-842-364-3 (1-366) x US-09-750-580-2 (1-1879)

QY 1 MetAlaSerMetAlaAlaValIleuThrTrpAlaLeuAlaLeuSerAlaIleuSerAla
Db 21 ATGGCAAGCATGGCTGCGCTGCTCAGCTGCTGCTCTTTCAGCGTTTCGGCC 80
QY 21 ThrGlnAlaArgLysGlyPheTrpAspTrpPheSerGlnThrSerGlyAspLysGlyArg 40
Db 81 ACCCAGGCGCGAAAGGCTTCTGGGACTTCTTCCAGCCAGCCAGCGCGGCGACAAAGCGAGC 140
QY 41 ValGlnGlnIleHisGlnGlnIleMetAlaArgLysProAlaThrLeuLysAspSerLeu 60
Db 141 GTGGAGCGATCCATCCACAGAGATGGCTCGCGAGCCCGCCAGCCCTTGAAGACAGCTT 200
QY 61 GluGlnAspLeuAsnAsnMetAsnLysPheLeuGlnLysLeuArgProLeuSerGlySer 80
Db 201 GAGCAGACCTCAACATATGACAAAGTCTCTGGAAGAGCTGAGGAGGCTCTGAGTGGAGC 260
QY 81 GluAlaProArgLeuProGlnAspProValGlyMetArgArgGlnLeuGlnGlnLeu 100
Db 261 GAGGCTCTCGGCTGCCACAGAGAGCCGGTGGCATGGCGGCGGCGGCTCCAGAGAGCTTG 320
QY 101 GluGlnValLysAlaArgLeuGlnProTrpMetAlaGlnAlaHisGlnLeuValGlyTrp 120
Db 321 GAGAGGTGAAGGCTCGGCTCCAGCCCTACATGGCAGAGCGGCGGAGCTGGTGGCTCG 380
QY 121 AsnLeuGlnGlyLeuArgGlnGlnLeuLysProTrpThrMetAspLeuMetGlnVal 140
Db 381 AATTGGAGGGCTGGCGGACGACCTGAAGCCCTACAGATGATGATGAGAGAGGTG 440
QY 141 AlaLeuArgValGlnGlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 441 GCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 500
QY 161 GlnLeuLeuGlnGlyValAspGlnAlaTrpAlaLeuLeuGlnGlnGlnGlnGlnGlnGln 180
Db 501 CAGTTGCTGGGGGCGGCTGGAGCAGCGCTTGCGCTTGTGCGAGGAGCTGCGAGCGCGCTG 560
QY 181 ValHisHisThrGlyArgPheLysGlnLeuPheHisProTrpAlaGlnSerLeuValSer 200
Db 561 GTGACACACAGCGGCGGCTTAAAGAGCTTCCACCCCTACGCGGAGAGGCTGTGTGAGC 620
QY 201 GilylIleGlyArgHisValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
Db 621 GCGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 680
QY 221 ProAlaArgLeuSerArgCysValGlnValLeuSerArgLysLeuThrLeuLysAlaLys 240
Db 681 CCCCGCGCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 740
QY 241 AlaLeuHisAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
Db 741 GCGCTGACGAGCGATCCAGAGAACCTGGACGAGCTGCGGAGAGAGCTGACAGAGAGCC 800

```

```

QY 261 PheAlaGlyThrGlyThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 801 TTTCGAGGCACTGGGAGCTAGAGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 860
QY 281 ValArgGlnArgLeuGlnIleAlaPheArgLysAspThrTrpLeuGlnIleAlaAlaPheTrp 300
Db 861 GTGGCGGCGGAGCTTCAAGGCTTTCGCGCAGGAGACCTTACCTGCGATAGCTGCTTCACT 920
QY 301 ArgAlaIleAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
Db 921 CGGCGCATGACAGAGAGAGCTAGAGAGTCCAGACAGCTGCGGCGGCGGCGGCGGCGGCGG 980
QY 321 HisSerAlaPheAlaProGlnPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 340
Db 981 CACAGTGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1040
QY 341 GlnAlaArgLeuAspAspLeuTrpGlnAspIleThrHisSerLeuHisAspGlnGlnHis 360
Db 1041 CAGGCGGCTGTGATGATCTGCTGGGAGACATCTACTACAGGCTTCTATGACAGGCGCAC 1100
QY 361 SerHisLeuGlyAspPro 366
Db 1101 AGCCATCTGGGGAGCCCC 1118

RESULT 2
US-09-750-580-4
: Sequence 4, Application US/09750580
: Patent No. 6455280
: GENERAL INFORMATION:
: APPLICANT: Yen, Frances
: APPLICANT: Denison, Blake
: APPLICANT: Bour, Barbara
: APPLICANT: Bihain, Bernard
: APPLICANT: Dumas Milne Edwards, Jean-Baptiste
: APPLICANT: Duclert, Aymeric
: APPLICANT: Bouquelier, Lydie
: APPLICANT: Ebbels-Reed, Dana
: APPLICANT: Salter-Cid, Luisa
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
: FILE REFERENCE: 89,052,CIP
: CURRENT FILING DATE: 2000-12-28
: PRIOR FILING DATE: 2000-12-28
: PRIOR FILING DATE: 2000-06-21
: PRIOR FILING DATE: 2000-06-21
: PRIOR APPLICATION NUMBER: PCT/IB00/0101
: PRIOR FILING DATE: 2000-06-21
: PRIOR APPLICATION NUMBER: PCT/IB99/02058
: PRIOR FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: US 49/469/099
: PRIOR FILING DATE: 1999-12-21
: PRIOR APPLICATION NUMBER: US 60/113,686
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: US 60/141,032
: PRIOR FILING DATE: 1999-06-25
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patent.pm
: SEQ ID NO 4
: LENGTH: 5381
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..918
: OTHER INFORMATION: 5' regulatory region
: NAME/KEY: exon
: LOCATION: 919..930
: OTHER INFORMATION: exon 1
: NAME/KEY: exon
: LOCATION: 1442..1498
: OTHER INFORMATION: exon 2
: NAME/KEY: exon
: LOCATION: 1613..1724

```

```

OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 2243..3940
OTHER INFORMATION: exon 4
NAME/KEY: misc_feature
LOCATION: 3941..5381
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 319
OTHER INFORMATION: 17-42-319 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 3213
OTHER INFORMATION: 17-41-250 : polymorphic base C or T
NAME/KEY: conflict
LOCATION: 1241
OTHER INFORMATION: 17-39-343 : T in ref genbank AC007707
NAME/KEY: conflict
LOCATION: 1447
OTHER INFORMATION: 17-40-202 : G in ref genbank AC007707
NAME/KEY: primer_bind
LOCATION: 1..11022
OTHER INFORMATION: 17-42..rp
NAME/KEY: primer_bind
LOCATION: 553..11575
OTHER INFORMATION: 17-42..rp complement
NAME/KEY: primer_bind
LOCATION: 899..11920
OTHER INFORMATION: 17-39..pu
NAME/KEY: primer_bind
LOCATION: 1246..12267
OTHER INFORMATION: 17-40..pu
NAME/KEY: primer_bind
LOCATION: 1441..12461
OTHER INFORMATION: 17-39..rp complement
NAME/KEY: primer_bind
LOCATION: 1632..12651
OTHER INFORMATION: 17-40..rp complement
NAME/KEY: primer_bind
LOCATION: 2964..13984
OTHER INFORMATION: 17-41..pu
NAME/KEY: primer_bind
LOCATION: 3432..14454
OTHER INFORMATION: 17-41..rp complement
NAME/KEY: primer_bind
LOCATION: 300..318
OTHER INFORMATION: 17-42-319..mis
NAME/KEY: primer_bind
LOCATION: 320..338
OTHER INFORMATION: 17-42-319..mis complement
NAME/KEY: primer_bind
LOCATION: 3194..3212
OTHER INFORMATION: 17-41-250..mis
NAME/KEY: primer_bind
LOCATION: 3214..3232
OTHER INFORMATION: 17-41-250..mis complement
NAME/KEY: misc_binding
LOCATION: 307..331
OTHER INFORMATION: 17-42-319..probe
NAME/KEY: misc_binding
LOCATION: 3201..3225
OTHER INFORMATION: 17-41-250..probe
US-09-750-580-4

```

Alignment Scores:

```

Pred. No.: 4.3e-305      Length: 5381
Score: 312.00           Matches: 312
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 85.25%             Indels: 0
DB: 4                     Gaps: 0

```

US-09-842-364-3 (1-366) x US-09-750-580-4 (1-5381)

```

Oy 55 ThrleuLysAspSerLeuGlnAspLeuAsnMeLanLysPheLeuGluLysLeu 74
Db 2244 ACCCTGAAGACAGCCTTGACACAGACTCAACAAATATGAAACAGTTCTTGAAAAGCTG 2303
Oy 75 ArgProLeuSerGlySerGlnAlaProArgLeuProGlnAspProValGlyMetLArg 94
Db 2304 AGGCCTTGAGTGGAGAGGAGGCTCTGCGTCCACAGACAGCCCGTGGATCGCGCGC 2363
Oy 95 GlnLeuGlnGluGluLeuGluGluValLysAlaArgLeuGlnProTyrMetAlaGlnAla 114
Db 2364 CAGCTGACGAGGAGATTGGAGAGAGTGAAGGCTCCCTCCAGCCCTCATGACGAGAGCG 2423
Oy 115 HisGluLeuValGlyTyrPasnLeuGlnGlyLeuArgGlnGlnLeuLysProTyrThiMet 134
Db 2424 CACGAGCTGGTGGCTGGAAATTTGAGGGCTTGGGGACGAACTGAAACCTTACACGATG 2483
Oy 135 AspLeuMetGluGlnValAlaLeuArgValGlnGluLeuGlnGlnValAla 154
Db 2484 GATCTGATGAGACAGACAGTGGCCCTGCGGCTCCAGACAGCTGCAGACAGAGTTCGCGTGC 2543
Oy 155 GlyLysAspThrLysAlaGlnLeuLeuGlyGlyValAspGluAlaTrrAlaLeuLeuGln 174
Db 2544 GGGAGAGACACACAGGCCCACTTCTGGGGGCGTGGACGAGGCTTGGCTTCTGCTGAG 2603
Oy 175 GlyLeuGlnSerArgValAlaHisHisThrGlyArgPheLysGluLeuPheHisProTyr 194
Db 2604 GGACTGCAGAGCCCGGTGGTCCACACACAGCCGCTTCAAGAGCTTCTCCACCCATAC 2663
Oy 195 AlaGluSerLeuValSerGlyIleGlyArgHisValGlnGluLeuHisArgSerValAla 214
Db 2664 GCCGAGACCCCTGTCAGCGGATCGGCGCCACGACGACGAGCTGCAGCCGACGCTGGCT 2723
Oy 215 ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys 234
Db 2724 CCGCAGCCGCCCGCCAGCCCGCCGCGCTCAGTGTGCTGTGCAAGGTCTTCCCGAAG 2783
Oy 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnAsnLeuAspGlnLeuArg 254
Db 2784 CTCAGCGCTCAAGGCCAAGGCCCTGACAGCAGCATCCAGCAAGAACTTGAGCAGCTGCGC 2843
Oy 255 GluGluLeuSerAlaAlaPheAlaGlyThrGlyThrGluGlnAlaGlyProAspPro 274
Db 2844 GAAGAGCTTCACAGACAGCTTTCAGCCACTGCGACTGAGAAAGGCCCGCCCGACCC 2903
Oy 275 GluMetLeuSerGluGluValArgGlnArgGlnArgGlnAlaPheArgGlnAspThrLysLeu 294
Db 2904 CAGATGCTCTCCGAGGAGGTGCGCCAGCAGCTTCAGGCTTTCGCCGAGACCTACTG 2963
Oy 295 GlnIleAlaIaPheThrArgAlaIleAspGlnGlnArgGlnGlnGlnGlnGlnGlnGln 314
Db 2964 CAAATAGCTGCTTCACTCCGCGCATGACAGAGACTGAGAGGTCCAGCAGCAGCTG 3023
Oy 315 AlaProProProProGlnHisSerAlaPheAlaProGluPheGlnGlnInThrAspSerGly 334
Db 3024 CCCCACCTCCACACAGCCACAGCTTCGCCCAAGTTTCAACAAACAGACAGCTGCG 3083
Oy 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTrpGluAspIleThrHisSer 354
Db 3084 AAGGTTCTGAGCAAGCTGCGAGGCCGCTGTGATACCTGTGGGAAGACATCACTACAGC 3143
Oy 355 LeuHisAspGlnGlnHisSerHisLeuGlyAspPro 366
Db 3144 CTTGATGACGAGGCCACAGCCTCTGCGGAGCC 3179

```

RESULT 3

US-09-750-580-1
Sequence 1, Application US/09750580
Patent No. 6455280

GENERAL INFORMATION:

APPLICANT: Yen, Frances
APPLICANT: Denison, Blake
APPLICANT: Bour, Barbara
APPLICANT: Bihain, Bernard


```

: LOCATION: 42206-.42230
: OTHER INFORMATION: 20-841-149 .probe
: NAME/KEY: misc.binding
: LOCATION: 45430-.45454
: OTHER INFORMATION: 20-842-115 .probe
: NAME/KEY: misc.binding
: LOCATION: 77046-.77070
: OTHER INFORMATION: 20-853-415 .probe
: LS-09-750-580-1

```

Alignment Scores:	
pred. No.:	6.45e-304
Score:	312.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	85.25%
DB:	4
Gaps:	0
Length:	81000
Matches:	312
Conservative:	0
Mismatches:	0
Indels:	0

US-09-842-364-3 (1-366) x US-09-750-580-1 (1-81001)

QY	55	ThrLeuLysASpSerLeuGluGlnAspLeuAsnMetLeuLysPheLeuLulysLeu	74
Db	14272	ACCCTGAAMAGACGCTTGAGCAACACTCAACATATATGAACAGTCTCTGGAAAAAGCTG	14333
QY	75	ArgProLeuSerGlySerClnAlaProArgLeuProGlnAspProValGlyMetArgArg	94
Db	14332	AGGCTCTGAGTGGGAGGAGGAGCTCTCCGCGCTCCACAGAGACCCGGTGGGCTGCGGCGG	14391
QY	95	GlnLeuGlnGlnGluLeuGlnGluValLysAlaArgLeuGlnProTyrMetAlaGluAla	114
Db	14382	CAGCTGCAGGAGAGACTGCAGAGCTGAAGGCTCCCTCCACGCTTACCTGGCAGAGCGCC	14451
QY	115	HisGluLeuValGlyTyrAsnLeuLeuGlyLeuMetArgGlnGlnLeuLysProTyrThrMet	134
Db	14452	CACGAGCTGGTGGGCTGGAAATTGGAGGGCTTGCGGCGACAGCACTGAAGCCCTACACGATG	14511
QY	135	AspLeuMetClnGlnValAlaLeuArgValGlnGluLeuGlnGlnGluLeuArgValVal	154
Db	14512	GATCTGATGCGAGACAGAGTGGCTCGCGCTGCGAGAGCTGCACAGACAGACTGGCGCGGTG	14571
QY	155	GlyGluAspPThrLysAlaGluLeuLeuGlnGlyValAlaArgLulAlaTrpAlaLeuGln	174
Db	14572	GCGGAGACACCCAAAGGCCACAGTCTGTGGCGGCGCTGCAGACAGCCTTGGCTTGTGCGAG	14631
QY	175	GlyLeuGlnSerArgValValHisHisThrGlyArgPheLysGluLeuPheHisProTyr	194
Db	14632	GGACTGCAGAGCGCGGTGGTGTGCACACACAGCGCGCTTCAAGAGCTCTCCACCATATAC	14691
QY	195	AlaGluLysLeuValSerGlyIleGlyArgHisValGlnGluLeuHisArgSerValAla	214
Db	14692	GCCGAGACCTGGTAGGCGGATCGGGCGCCACGTGCAGAGACTGCACCGCATGTGGCT	14751
QY	215	ProHisAlaProAlaSerProAlaArgLeuSerArgGlySvalGlnValLeuSerArgLys	234
Db	14752	CCGCAAGCCCCCGCAGCCGCCCGCGCTCACTCCCTGCTGCAGAGTGTCTCCCGAAG	14811
QY	235	LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnAsnLeuAspGlnLeuArg	254
Db	14812	CTCACGGCTCAAGGCGCAAGGCCCTGCAGCAGCAGCATCCAGAGAACCTGGACAGCTGGCC	14871
QY	255	GlnGluLeuSerArgAlaPheAlaGlyThrGlyThrGlnGluGluValAlaGlyProAspPro	274
Db	14872	GAAAGGCTCACAGAGCTTTGTCAGGCACTGGGACTGGAAGAGGGCGCGCCGGAACCCC	14931
QY	275	GlnMetLeuSerGlnGluValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeu	294
Db	14932	CAGATGCTCTCCGAGAGAGTGGCCACGCACTTCCAGGCTTTCCGCGACAGACACCTTACCTG	14991
QY	295	GlnIleAlaAlaPheThrArgAlaAlaLeuAspGlnGlnThrGlnGluValAlaGlnGlnLeu	314
Db	14992	CAGATAGCTGCTTCACTACGCGCCATCGAACACAGAGACTGAAGAGAGTCCACAGACAGCTG	15051
QY	315	AlaProProProProGlnHisSerAlaPheAlaProGluPheGlnGlnThrAspSerGly	334

Accession	Protein	Length
Db 15052	GGCGCCACCTCCACACGGCCACAGCTCCTTGGCCCGACAGTTPPCCAAACAAACACAGCTGC	15111
QY 335	LysValIleuSerLysIleuGlnIaIaArgIeuAspAspIleuTrGluAspIleuThrIleuSer	354
Db 15112	AAGCTTCTGACCAACCTCAGGCCCGCTCTGCATGACCTCTGGGAAGACATCCACACAC	15177
QY 355	IeuHtIAspGlnGlyHisSerHisIleuGlyAspPro	366
Db 15172	CTTCATGACCAAGGCCACACAGCATTTGGGGGACCCC	15207

QY	355	LeuHisAspGlnGlyHisSerHisLeuGlyAspPro	366
Db	15172	CTTCATGACCAAGGCGACAGCCATCTGGGGGACCC	15207

RESULT 4

US-09-276-531-115
; Sequence 115, Application US/09276531
; Patent No. 6183068

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Yuc, Henry
APPLICANT: Reddy Boona

APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah P.

5 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
6 RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION

CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 134

ADDRESSEE: INCYTE PHARMAC
STREET. 3174 PORTER DRIVE

CITY: PALO ALTO
STATE: CALIFORNIA

COUNTRY: USA
ZTP. 9430A

COMPUTER READABLE
MEDIUM TYPE.

```
COMPUSER:  IH
OPBPAT"INC  sys
```

```

; SOFTWARE: word perfect
CURRENT APPLICATION DATA

```

APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith

CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLIICATION NUMBER: 60/079,677
FILING DATE: March 27 1998

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION.

NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,911

REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:

LEUOH. 332 base pair
TYPE: nucleic acid

```

; STANDEDNESS: 51
; TOPOLOGY: linear
;

```

LIBRARY: LIVRUTO

Program Location:	
Pred. No.:	9,9e-169
Score:	176.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	48.09%
DB:	4
Length:	592
Matches:	176
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-842-364-3 (1-366) x US-09-276-531-115 (1-592)

```

Qy 1 MetAlaSerMetAlaAlaValLeuThrTrpAlaLeuAlaLeuSerAlaPheSerAla 20
    |||||||
Db 35 ATGGCAACGATGGCTGCCCTGCTCAGCTGGCTTCTTTTACGCTTTTCGGCC 94
Qy 21 ThrGlnAlaArgLysGlyPheTrpAspTyrPheSerGlnThrSerGlyAspLysGlyArg 40
    |||||||
Db 95 ACCGAGCAGCAGAAAGGCTTCTGGAGTACTTCAAGCAGACAGCGGGAGCAAGAGGAGG 154
Qy 41 ValGlnGlnIleHISGlnGlnLysMetAlaArgGluProAlaThrLeuLysAspSerLeu 60
    |||||||
Db 155 GTGAGCAGATCCATCGCAGAAAGATGCTGCGAGCCCGGACCTGAAAGACAGCCTT 214
Qy 61 GluGlnAspLeuAsnAsnMetAsnLysPheLeuGluLysLeuArgProLeuSerGlySer 80
    |||||||
Db 215 GAGCAAGACCTCAACATATGCAACAACCTTCTGAAAGAGCTGAGCCCTTCAGTGGAGCC 274
Qy 81 GluAlaProArgLeuProGlnAspProValGlyMetArgArgGlnLeuGlnGluLeu 100
    |||||||
Db 275 GAGGCTCCTCGGCTCCACAGACCCCGTGCGCATGCGCGGAGCTGCAGAGAGAGATTG 334
Qy 101 GluGlnValAlaLysAlaArgLeuGlnProTyrMetAlaGlnAlaHISGlnLeuValGlyTyr 120
    |||||||
Db 335 GAGGAGGTGAAGGCTCGCTCCAGCCCTACATGGCAGAGCGCAGAGCTGGTGGCTGG 394
Qy 121 AsnLeuGlnGlyLeuArgGlnGlnLeuLysProTyrThrMetAspLeuMetGlnGlnVal 140
    |||||||
Db 395 AATTTGAGCGCTTCCGCGCAGCACTGACGCCATGACATGATCGATCGAGCAGCAGCTG 454
Qy 141 AlaLeuArgValGlnGlnLeuGlnGlnGlnGlnLeuArgValAlaGlyLysProThrLysAla 160
    |||||||
Db 455 GCCCTGGCTGAGGAGGCTGAGGAGCAGGACATTCGCTGTGGGGAGACACCAAGGCC 514
Qy 161 GlnLeuLeuGlnGlyValAlaAspGlnAlaTrpAlaLeuLeuGlnGlnLeu 176
    |||||||
Db 515 CAGTGTGGGGGGGCGTGGAGCAGGCTTGGGCTTGTGTCAGAGGAGCTG 562

RESULT 5
US-08-487-811A-23
; Sequence 23, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Cleveli, Olivier
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,811A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5883226nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; Alignment Scores:
; Pred. No.: 3.43
; Score: 8.00
; Percent Similarity: 100.00%
; Best Local Similarity: 100.00%
; Query Match: 2.19%
; DB: 2
; Gaps: 0
US-09-842-364-3 (1-366) x US-09-060-694-23 (1-48)
Qy 82 AlaProArgLeuProGlnAspPro 89
    |||||||
Db 5 GCCGCCGCGCTCCCGCAGAGCCCGC 28

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-487-811A-23
; Alignment Scores:
; Pred. No.: 3.43
; Score: 8.00
; Percent Similarity: 100.00%
; Best Local Similarity: 100.00%
; Query Match: 2.19%
; DB: 2
; Gaps: 0
US-09-842-364-3 (1-366) x US-08-487-811A-23 (1-48)
Qy 82 AlaProArgLeuProGlnAspPro 89
    |||||||
Db 5 GCCGCCGCGCTCCCGCAGAGCCCGC 28

RESULT 6
US-09-060-694-23
; Sequence 23, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Cleveli, Olivier
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-MM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; Alignment Scores:
; Pred. No.: 3.43
; Score: 8.00
; Percent Similarity: 100.00%
; Best Local Similarity: 100.00%
; Query Match: 2.19%
; DB: 4
; Gaps: 0
US-09-842-364-3 (1-366) x US-09-060-694-23 (1-48)
Qy 82 AlaProArgLeuProGlnAspPro 89

```



```

|||||
DB      5 GCGCCCGCCTCCCGCAGACCCC 28

RESULT 7
US-08-623-471-6
; Sequence 6, Application US/08623471
; Patent No. 5846823
; GENERAL INFORMATION:
; APPLICANT: Allelix Biopharmaceuticals Inc
; APPLICANT: Owolabi, Joshua
; APPLICANT: Rampersad, Vikarna
; APPLICANT: Ramboj, Rajender
; TITLE OF INVENTION: STABLE D4 CELL LINES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allelix Biopharmaceuticals Inc
; STREET: 6850 Goreway Drive
; CITY: Mississauga
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/623,471
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA94/00538
; FILING DATE: 27-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RIDOUT & MAYBEE, Attn. Robert G. Hiron
; REFERENCE/DOCKET NUMBER: ALLEL/51B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416)-868-1482
; TELEFAX: (416)-362-0823
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-623-471-6

Alignment Scores:
Pred. No.: 7.49      Length: 105
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.19%      Indels: 0
                        Gaps: 0
DB: 2

US-09-842-364-3 (1-366) x US-08-623-471-6 (1-105)

OY      82 AlaprogLeuprogInaspro 89
DB      10 GCGCCCGCCTCCCGCAGACCCC 33

RESULT 8
US-08-623-471-7
; Sequence 7, Application US/08623471
; Patent No. 5846823
; GENERAL INFORMATION:
; APPLICANT: Allelix Biopharmaceuticals Inc
; APPLICANT: Owolabi, Joshua
; APPLICANT: Rampersad, Vikarna
; APPLICANT: Ramboj, Rajender
; TITLE OF INVENTION: STABLE D4 CELL LINES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allelix Biopharmaceuticals Inc

```

```

; STREET: 6850 Goreway Drive
; CITY: Mississauga
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/623,471
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA94/00538
; FILING DATE: 27-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RIDOUT & MAYBEE, Attn. Robert G. Hiron
; REFERENCE/DOCKET NUMBER: ALLEL/51B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416)-868-1482
; TELEFAX: (416)-362-0823
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-623-471-7

Alignment Scores:
Pred. No.: 14.3     Length: 201
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.19%      Indels: 0
                        Gaps: 0
DB: 2

US-09-842-364-3 (1-366) x US-08-623-471-7 (1-201)

OY      82 AlaprogLeuprogInaspro 89
DB      10 GCGCCCGCCTCCCGCAGACCCC 33

RESULT 9
US-08-623-471-8
; Sequence 8, Application US/08623471
; Patent No. 5846823
; GENERAL INFORMATION:
; APPLICANT: Allelix Biopharmaceuticals Inc
; APPLICANT: Owolabi, Joshua
; APPLICANT: Rampersad, Vikarna
; APPLICANT: Ramboj, Rajender
; TITLE OF INVENTION: STABLE D4 CELL LINES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allelix Biopharmaceuticals Inc
; STREET: 6850 Goreway Drive
; CITY: Mississauga
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/623,471
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA94/00538
; FILING DATE: 27-SEP-1994
; ATTORNEY/AGENT INFORMATION:

```



```

|||||
DB      350 GCGCCCGCTCCCGCAGAGACCC 373

RESULT 12
US-08-333-977-8
; Sequence 8, Application US/08333977
; Patent No. 5594108
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civeilli, Olivier
; TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Wilcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,977
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,618
; FILING DATE: 7 DEC 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5594108man, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1234
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1..409
; FEATUERE:
; NAME/KEY: CDS
; LOCATION: 2..409
; US-08-333-977-8

Alignment Scores:
Pred. No.: 29.1 Length: 409
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.19% Indels: 0
Gaps: 0

US-09-842-364-3 (1-366) x US-08-333-977-8 (1-409)
OY      82 AlaProArgLeuProGlnAspPro 89
DB      350 GCGCCCGCTCCCGCAGAGACCC 373

RESULT 13
US-07-928-611-8
; Sequence 8, Application US/07928611
; Patent No. 5569601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civeilli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Wilcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,611
; FILING DATE: 19920810
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5569601man, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 563 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1..563
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /standard_name= "Alternate Exon 3: D4.2"
; OTHER INFORMATION: /note= "This sequence represent the sequence of
; the third exon of allele D4.2 of the human D4
; dopamine receptor gene"
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 257..262
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "Polymorphic PstI site"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /label= PstI
; OTHER INFORMATION: /note= "this feature is the site of one of the
; restriction enzymes whereby digestion of genomic
; DNA produces a RFLP "
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 346..442
; OTHER INFORMATION: /rpt_type= "random"
; OTHER INFORMATION: /rpt_unit= 348 .. 396
; OTHER INFORMATION: /note= "This sequence represents one of 7 known
; alleles of human D4 dopamine receptor gene
; encoding a 16 amino acid sequence repeated twice"
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..563
; US-07-928-611-8

Alignment Scores:
Pred. No.: 40 Length: 563
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.19% Indels: 0

```

DB: 1 Gaps: 0

US-09-842-364-3 (1-366) x US-07-928-611-8 (1-563)

OY 82 AlaprogLeuprogInaspro 89
 |||
 Db 350 GCGCCCGCTCCCGAGACCC 373

RESULT 14

US-08-487-811A-8
 ; Sequence 8, Application US/08487811A
 ; Patent No. 5883226
 ; GENERAL INFORMATION:
 ; APPLICANT: Civeili, Olivier
 ; APPLICANT: Van Tol, Hubert H.M.
 ; TITLE OF INVENTION: A No. 5883226e1 Human Dopamine Receptor and Uses
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 ; STREET: 300 South Wacker Drive
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/487,811A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5883226nan, Kevin E
 ; REGISTRATION NUMBER: 35,303
 ; REFERENCE/DOCKET NUMBER: 90,1092-L
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-913-0001
 ; TELEFAX: 312-913-0002
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 563 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: exon
 ; LOCATION: 1..563
 ; IDENTIFICATION METHOD: experimental
 ; OTHER INFORMATION: /evidence= EXPERIMENTAL
 ; OTHER INFORMATION: /standard_name= "Alternate Exon 3: D4.2"
 ; OTHER INFORMATION: /note= "This sequence represent the sequence of
 ; OTHER INFORMATION: the third exon of allele D4.2 of the human D4
 ; OTHER INFORMATION: dopamine receptor gene"
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: 257..262
 ; IDENTIFICATION METHOD: experimental
 ; OTHER INFORMATION: /function= "Polymorphic PstI site"
 ; OTHER INFORMATION: /evidence= EXPERIMENTAL
 ; OTHER INFORMATION: /label= PstI
 ; OTHER INFORMATION: /note= "This feature is the site of one of the
 ; OTHER INFORMATION: restriction enzymes whereby digestion of genomic
 ; OTHER INFORMATION: DNA produces a RFLP "
 ; FEATURE:
 ; NAME/KEY: repeat_region
 ; LOCATION: 346..442
 ; OTHER INFORMATION: /rpt_type= "tandem"
 ; OTHER INFORMATION: /rpt_unit= 348..396
 ; OTHER INFORMATION: /note= "This sequence represents one of 7 known

OTHER INFORMATION: alleles of human D4 dopamine receptor gene
 ; OTHER INFORMATION: encoding a 16 amino acid sequence repeated twice
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2..563
 ; US-08-487-811A-8

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
8.00	100.00%	563	8
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	2.19%	Mismatches:	0
DB:	2	Indels:	0
		Gaps:	0

US-09-842-364-3 (1-366) x US-08-487-811A-8 (1-563)

OY 82 AlaprogLeuprogInaspro 89
 |||
 Db 350 GCGCCCGCTCCCGAGACCC 373

RESULT 15

US-09-060-694-8
 ; Sequence 8, Application US/09060694
 ; Patent No. 6203998
 ; GENERAL INFORMATION:
 ; APPLICANT: Civeili, Olivier
 ; APPLICANT: Van Tol, Hubert H.M.
 ; TITLE OF INVENTION: A No. 6203998e1 Human Dopamine Receptor and Uses
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 ; STREET: 300 South Wacker Drive
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/060,694
 ; FILING DATE: 15-APR-1998
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 6203998nan, Kevin E
 ; REGISTRATION NUMBER: 35,303
 ; REFERENCE/DOCKET NUMBER: 90,1092-XM
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-913-0001
 ; TELEFAX: 312-913-0002
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 563 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: exon
 ; LOCATION: 1..563
 ; IDENTIFICATION METHOD: experimental
 ; OTHER INFORMATION: /evidence= EXPERIMENTAL
 ; OTHER INFORMATION: /standard_name= "Alternate Exon 3: D4.2"
 ; OTHER INFORMATION: /note= "This sequence represent the sequence of
 ; OTHER INFORMATION: the third exon of allele D4.2 of the human D4
 ; OTHER INFORMATION: dopamine receptor gene"
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: 257..262

```
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "Polymorphic PstI site"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /label= PstI
OTHER INFORMATION: /note= "this feature is the site of one of the
OTHER INFORMATION: restriction enzymes whereby digestion of genomic
OTHER INFORMATION: DNA produces a RFLP "
FEATURE:
NAME/KEY: repeat_region
LOCATION: 346..442
OTHER INFORMATION: /rpt_type= "tandem"
OTHER INFORMATION: /rpt_unit= 348..396
OTHER INFORMATION: /note= "this sequence represents one of 7 known
OTHER INFORMATION: alleles of human D4 dopamine receptor gene
OTHER INFORMATION: encoding a 16 amino acid sequence repeated twice
FEATURE:
NAME/KEY: CDS
LOCATION: 2..563
US-09-060-694-8
```

```
Alignment Scores:
Pred. No.: 40
Score: 8.00 Length: 563
Percent Similarity: 100.00% Matches: 8
Best Local Similarity: 100.00% Conservative: 0
Query Match: 2.19% Mismatches: 0
DB: 4 Indels: 0
Gaps: 0
```

US-09-842-364-3 (1-366) x US-09-060-694-8 (1-563)

```
OY 82 AlaProArgLeuProGlnAspPro 89
    |||||||||||||||||||||
DB 350 GCGCCCGCGCCTCCCGCAGAGCCCC 373
```

Search completed: January 23, 2003, 19:24:03
Job time : 71 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 23, 2003, 17:48:17 : Search time 67 Seconds
(without alignments)
2454.193 Million cell updates/sec

Title: US-09-842-364-3
Perfect score: 366
Sequence: 1 NASMAVLTWALALSAFSA.....LMEDITHSLHDOGHSLGDP 366

Scoring table: OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 396772 seqs, 224632407 residues

Word size: 6

Total number of hits satisfying chosen parameters: 6466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-O/cg2n2.1/US07.spool/US09842364/runat.16012003.164830.10683/app.query.fasta.1.519
-DB-Published.Applications.NA -QWMT-lastap -SUFFIX-olig.rnpb -MINMATCH-0.1
-LOOPCL-0 -LOOPEXT-0 -ONITS-bits -START-1 -END-1 -MATRIX-oligo
-TRANS-human40.ccl -LIST-45 -DOCALIGN-200 -THR-SCORE-quality -THR-MIN-6
-ALIGN-15 -MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZ-500 -MINLEN-0
-MAXLEN-200000000 -USER-US09842364 -CGCN.1.1.33 -runat.16012003.164830.10683
-NCPU-6 -ICPU-3 -NO.XLPXY -NO.MMAP -LARGEIOBURY -NEG.SCORES-0 -WAIT -LONGIOG
-DEV.TIMEOUT-120 -WARN.TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6
-FGAPEXT-7 -FGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database: Published.Applications.NA.*

1: /cg2n2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2n2.6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cg2n2.6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2n2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cg2n2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
6: /cg2n2.6/ptodata/1/pubpna/PCR_US_PUBCOMB.seq:*
7: /cg2n2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2n2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2n2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cg2n2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cg2n2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cg2n2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cg2n2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cg2n2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	366	100.0	1858 10	US-09-835-996A-1
2	366	100.0	1877 10	US-09-835-996A-10
3	366	100.0	1879 10	US-09-751-877-2
4	312	85.2	5381 10	US-09-751-877-4

5	312	85.2	81001 10	US-09-751-877-1	Sequence 1, Appl1
6	279	76.2	1393 10	US-09-800-729-40	Sequence 40, Appl1
7	213	58.2	1334 10	US-09-349-015-1	Sequence 1, Appl1
8	70	19.1	468 10	US-09-800-729-209	Sequence 209, Appl
9	33	9.0	331 10	US-09-800-729-210	Sequence 210, Appl
10	9	2.5	301 10	US-09-983-965-4574	Sequence 4574, Ap
11	9	2.5	397 10	US-09-983-965-4360	Sequence 4360, Ap
12	12	2.5	405 10	US-09-960-352-10825	Sequence 10825, A
13	13	2.5	419 10	US-09-960-352-8211	Sequence 8211, Ap
14	14	2.5	832 9	US-09-938-824-2586	Sequence 2586, Ap
15	9	2.5	2850 10	US-09-815-842-7104	Sequence 7404, Ap
16	8	2.2	158 9	US-09-796-692-3304	Sequence 3304, Ap
17	8	2.2	436 10	US-09-983-965-4711	Sequence 4711, Ap
18	8	2.2	368 10	US-09-983-965-5038	Sequence 5038, Ap
19	8	2.2	445 10	US-09-864-761-2633	Sequence 2633, Ap
20	8	2.2	505 10	US-09-783-590-2423	Sequence 2423, Ap
21	8	2.2	522 10	US-09-864-761-32402	Sequence 32402, A
22	8	2.2	582 10	US-09-864-761-15898	Sequence 15898, A
23	8	2.2	687 10	US-09-735-705-38	Sequence 38, Appl
24	8	2.2	687 10	US-09-850-716A-38	Sequence 38, Appl
25	8	2.2	687 10	US-09-897-778-38	Sequence 38, Appl
26	5	2.2	755 10	US-09-809-545A-9	Sequence 9, Appl1
27	8	2.2	948 10	US-09-886-055-176	Sequence 176, App
28	8	2.2	1006 9	US-09-946-807-104	Sequence 104, App
29	8	2.2	1006 10	US-09-795-668-104	Sequence 104, App
30	8	2.2	1006 10	US-09-795-668-104	Sequence 104, App
31	8	2.2	1413 9	US-09-938-824-422	Sequence 422, App
32	8	2.2	1458 10	US-09-741-669-106	Sequence 106, App
33	8	2.2	1532 9	US-09-992-598-150	Sequence 150, App
34	8	2.2	1532 9	US-09-989-293A-150	Sequence 150, App
35	8	2.2	1532 9	US-09-989-735-150	Sequence 150, App
36	8	2.2	1532 9	US-09-990-44A-150	Sequence 150, App
37	8	2.2	1532 9	US-09-989-730-150	Sequence 150, App
38	8	2.2	1532 9	US-09-990-436-150	Sequence 150, App
39	8	2.2	1532 9	US-09-991-181-150	Sequence 150, App
40	8	2.2	1532 9	US-09-993-687-150	Sequence 150, App
41	8	2.2	1532 9	US-09-989-735A-150	Sequence 150, App
42	8	2.2	1532 9	US-09-997-653-150	Sequence 150, App
43	8	2.2	1532 9	US-09-997-653-150	Sequence 150, App
44	8	2.2	1532 9	US-10-174-590-121	Sequence 121, App
45	8	2.2	1532 9	US-10-175-737-121	Sequence 121, App

ALIGNMENTS

RESULT 1
US-09-835-996A-1
Sequence 1, Application US/09835996A
Patent No. US20020142953A1
GENERAL INFORMATION:
APPLICANT: Baillinger, Dennis
APPLICANT: Loeb, Debra
APPLICANT: Montgomery, Julie
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing
APPLICANT: Wehrman, Tom
APPLICANT: Drmanac, Radoje
APPLICANT: Ren, Feiyan
APPLICANT: Qian, Xiaohong
APPLICANT: Wang, Durrul
TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
FILE REFERENCE: 28110/35915A
CURRENT APPLICATION NUMBER: US/09/835, 996A
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: US 60/197,137
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: US 09/714,936
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: US 09/667,298


```

Db 201 GAGCAAGACCTCAACATATGACAGTCTCTGGAAAAAGCTGAGGCTCTGATGGAGC 260
QY 81 GUAAlaAroArgLeuProGlnAspProValGlyMetArgArgGlnLeuGlnGlnGln 100
Db 261 GAGCTCTCTCTGCTCCACAGAGACCGGTGGCGATGGCGGCGAGCTGCAGAGAGCTTG 320
QY 101 GUGUValAlaAlaArgLeuGlnProTyrMetAlaGlnAlaHisGlnLeuValGlyTyr 120
Db 321 GAGAGAGTGAAGCTCTGCTCCACAGCTTACATGCGACAGCGGCGAGCTCTGCTGCTGC 380
QY 121 AsnLeuGlnGlyLeuArgGlnGlnGlnLeuGlnGlnGlnGlnGlnGlnGlnGln 140
Db 381 AATTGGAGAGCTTGGCGAGCACTGAGACCTTACACAGATGATGATGATGATGATGATG 440
QY 141 AlaLeuArgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 441 GCGCTGCGCTGCGAGCGCTGCAGAGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 500
QY 161 GlnLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 501 CAGTTCCTGGCGGCGCTGCAGAGAGCTTGGCGCTTTCCTCCAGAGAGCTGCAGAGAGCT 560
QY 181 ValHisHisThrGlyArgPheLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 561 GTGACACACACCGCGCTTCAAGAGCTCTTCCACCCATACGCGAGAGCTGATGAGC 620
QY 201 GUYIIGLYArgHisValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
Db 621 GCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 680
QY 221 ProAlaArgLeuSerArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
Db 681 CCGCGCGCGCTCAAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 740
QY 241 AlaLeuHisAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
Db 741 GCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
QY 261 PheAlaGlyThrGlyThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 801 TTTGCGAGCGCTGCGAGTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860
QY 281 ValArgGlnArgLeuGlnAlaPheArgGlnAspThrTyrLeuGlnGlnGlnGlnGln 300
Db 861 CTGCGCGCGCGCTGCGAGCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 920
QY 301 ArgAlaAlaAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
Db 921 CCGGCGCGCTGCGAGTGAAGAGCTGCGAGAGCTGCGCGCGCGCGCGCGCGCGCGCG 980
QY 321 HisSerAlaPheAlaProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 340
Db 981 CACAGTTCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040
QY 341 GlnAlaArgLeuAspAspLeuTrpGlnAspIleThrHisSerLeuHisAspGlnGln 360
Db 1041 CAGCGCGCGCTGCGAGTGAAGAGCTGCGAGAGCTGCGAGAGCTGCGAGAGCTGCG 1100
QY 361 SerHisLeuGlnAspPro 366
Db 1101 AGCCATCTGGGAGACCC 1118

```

```

? APPLICANT: Bougueleret, Lydie
? APPLICANT: Ebbels-Reed, Dana
? APPLICANT: Salter-Cid, Luisa
? TITLE OR INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
? FILE REFERENCE: 89.US.1.RG
? CURRENT APPLICATION NUMBER: US/09/751,877
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: Patent.pm
? SEQ ID NO 4
? LENGTH: 5381
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc-feature
? LOCATION: 1..918
? OTHER INFORMATION: 5'regulatory region
? NAME/KEY: exon
? LOCATION: 919..930
? OTHER INFORMATION: exon 1
? NAME/KEY: exon
? LOCATION: 1442..1498
? OTHER INFORMATION: exon 2
? NAME/KEY: exon
? LOCATION: 1613..1724
? OTHER INFORMATION: exon 3
? NAME/KEY: exon
? LOCATION: 2243..3940
? OTHER INFORMATION: exon 4
? NAME/KEY: misc-feature
? LOCATION: 3941..5381
? OTHER INFORMATION: 3'regulatory region
? NAME/KEY: allele
? LOCATION: 319
? OTHER INFORMATION: 17-42-319 : polymorphic base C or T
? NAME/KEY: allele
? LOCATION: 3213
? OTHER INFORMATION: 17-41-250 : polymorphic base C or T
? NAME/KEY: conflict
? LOCATION: 1241
? OTHER INFORMATION: 17-39-343 : T in ref genbank AC007707
? NAME/KEY: conflict
? LOCATION: 1447
? OTHER INFORMATION: 17-40-202 : G in ref genbank AC007707
? NAME/KEY: primer_bind
? LOCATION: 1..11022
? OTHER INFORMATION: 17-42-pu
? NAME/KEY: primer_bind
? LOCATION: 553..11575
? OTHER INFORMATION: 17-42-rp complement
? NAME/KEY: primer_bind
? LOCATION: 899..11920
? OTHER INFORMATION: 17-39-pu
? NAME/KEY: primer_bind
? LOCATION: 1246..12267
? OTHER INFORMATION: 17-40-pu
? NAME/KEY: primer_bind
? LOCATION: 1441..12461
? OTHER INFORMATION: 17-39-rp complement
? NAME/KEY: primer_bind
? LOCATION: 1632..12651
? OTHER INFORMATION: 17-40-rp complement
? NAME/KEY: primer_bind
? LOCATION: 2964..13984
? OTHER INFORMATION: 17-41-pu
? NAME/KEY: primer_bind
? LOCATION: 3432..14454
? OTHER INFORMATION: 17-41-rp complement
? NAME/KEY: primer_bind
? LOCATION: 300..318
? OTHER INFORMATION: 17-42-319 mis
? NAME/KEY: primer_bind
? LOCATION: 320..338

```

OTHER INFORMATION: 17-42-319.mis complement
 NAME/KEY: primer_bind
 LOCATION: 3194..3212
 OTHER INFORMATION: 17-41-250.mis
 NAME/KEY: primer_bind
 LOCATION: 3214..3232
 OTHER INFORMATION: 17-41-250.mis complement
 NAME/KEY: misc_binding
 LOCATION: 307..331
 OTHER INFORMATION: 17-42-319.probe
 NAME/KEY: misc_binding
 LOCATION: 3201..3225
 OTHER INFORMATION: 17-41-250.probe
 US-09-751-877-4

Alignment Scores:
 Pred. No.: 7.62e-297 Length: 5381
 Score: 312.00 Matches: 312
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.25% Indels: 0
 DB: 10 Caps: 0

US-09-842-364-3 (1-366) x US-09-751-877-4 (1-5381)

QY 55 ThrLeuLysAspSerLeuGlnGlnAspLeuAsnMetLysPheLeuGluLysLeu 74
 DB 2244 ACCCTGAAACACACACCTTGAACCAAGACCTTGAACCAAGCTTCTGGAACACCTG 2303
 QY 75 ArgProLeuSerGlySerGlnAlaProArgLeuProGlnAspProValGlyMetLargArg 94
 DB 2304 AGCCCTCTGAGTGAGGAGGAGGCTCTCCGCTCCACAGACCGGAGGAGGAGGAGG 2363
 QY 95 GlnLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 114
 DB 2364 CAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2423
 QY 115 HisGlnLeuValGlyTTPAnLeuGlnGlyLeuArgGlnGlnGlnGlnGlnGlnGlnGln 134
 DB 2424 CAGCAGCTGCTGGGCTGGGATTTGCAAGCCCTTCGCCGACGACCTTACACCATG 2483
 QY 135 AspLeuMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 154
 DB 2484 GATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2543
 QY 155 GlyGlnAspThrLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 174
 DB 2544 GGGGAAAGACACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2603
 QY 175 GlyLeuGlnSerArgValAlaHisHisThrGlyArgPheLysGlnLeuPheHisProTyr 194
 DB 2604 GCACCTGCAGAGCCGCGCTGCGACACACCCGCCCTTCGCAAGACCTTCGACCCATAC 2663
 QY 195 AlaGlnSerLeuValSerGlyIleGlyArgHisValGlnGlnLeuHisAspSerValAla 214
 DB 2664 GCCGAGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2723
 QY 215 ProHisAlaProLysSerProAlaArgLysSerArgCysValGlnValLeuSerArgLys 234
 DB 2724 CCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2783
 QY 235 LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnGlnGlnGlnGlnGlnGln 254
 DB 2784 CTCACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2843
 QY 255 GlnGlnLeuSerArgAlaPheAlaGlyThrGlyThrGlnGlnGlnGlnGlnGlnGlnGln 274
 DB 2844 GAAGAGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2903
 QY 275 GlnMetLeuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 294
 DB 2904 CAGATGCTCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2963

QY 295 GlnIleAlaAlaPheThrArgAlaIleAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 314
 DB 2964 CAGATAGCTGCTTCTTACCTGCGGCGCCACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3023
 QY 315 AlaProProProProGlnHisSerAlaPheAlaProGlnGlnGlnGlnGlnGlnGlnGln 334
 DB 3024 GCGCAGCTCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3083
 QY 335 LysValLeuSerLysLeuGlnAlaArgLeuAspAspLeuTyrPGLuAspPheThrHisSer 354
 DB 3084 AAGCTTCGACCAAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3143
 QY 355 LeuHisAspGlnGlnHisSerHisLeuGlnAspPro 366
 DB 3144 CTCATCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3179

RESULTS 5
 US-09-751-877-1
 Sequence 1, Application US/09751877
 Patent No. US20020142949A1
 GENERAL INFORMATION:
 APPLICANT: Yen, Frances
 APPLICANT: Dentison, Blake
 APPLICANT: Bour, Barbara
 APPLICANT: Blhain, Bernard
 APPLICANT: Dumas Milne Edwards, Jean-Baptiste
 APPLICANT: Duclet, Aymeric
 APPLICANT: Bougueterec, Lydie
 APPLICANT: Ebbets-Keed, Dana
 APPLICANT: Salter-Cid, Luisa
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
 FILE REFERENCE: 89 US3 REG
 CURRENT APPLICATION NUMBER: US/09/751,877
 CURRENT FILING DATE: 2000-12-28
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentL.pm
 SEQ ID NO 1
 LENGTH: 81001
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 10946..12946
 OTHER INFORMATION: 5'regulatory region
 NAME/KEY: exon
 LOCATION: 12947..12958
 OTHER INFORMATION: exon 1
 NAME/KEY: exon
 LOCATION: 13470..13526
 OTHER INFORMATION: exon 2
 NAME/KEY: exon
 LOCATION: 13641..13752
 OTHER INFORMATION: exon 3
 NAME/KEY: exon
 LOCATION: 14271..15968
 OTHER INFORMATION: exon 4
 NAME/KEY: misc_feature
 LOCATION: 15969..17969
 OTHER INFORMATION: 3'regulatory region
 NAME/KEY: allele
 LOCATION: 1239
 OTHER INFORMATION: 20-828-311 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 12347
 OTHER INFORMATION: 17-42-319 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 15241
 OTHER INFORMATION: 17-41-250 : polymorphic base C or T
 NAME/KEY: allele
 LOCATION: 42218
 OTHER INFORMATION: 20-841-149 : polymorphic base A or G
 NAME/KEY: allele
 LOCATION: 45442

1	OTHER INFORMATION: 20-842-115	: polymorphic
2	NAME/KEY: allele	
3	LOCATION: 77058	
4	OTHER INFORMATION: 20-853-415	: polymorphic
5	NAME/KEY: primer_bind	
6	LOCATION: 929..949	
7	OTHER INFORMATION: 20-828..pu	
8	NAME/KEY: primer_bind	
9	LOCATION: 1357..1377	
10	OTHER INFORMATION: 20-828..rp	complement
11	NAME/KEY: primer_bind	
12	LOCATION: 14992..15012	
13	OTHER INFORMATION: 17-41..pu	
14	NAME/KEY: primer	
15	LOCATION: 15460..15482	
16	OTHER INFORMATION: 17-41..rp	complement
17	NAME/KEY: primer_bind	
18	LOCATION: 42070..42090	
19	OTHER INFORMATION: 20-841..pu	
20	NAME/KEY: primer_bind	
21	LOCATION: 42572..42591	
22	OTHER INFORMATION: 20-841..rp	complement
23	NAME/KEY: primer_bind	
24	LOCATION: 45328..45347	
25	OTHER INFORMATION: 20-842..pu	
26	NAME/KEY: primer_bind	
27	LOCATION: 45663..45683	
28	OTHER INFORMATION: 20-842..rp	complement
29	NAME/KEY: primer_bind	
30	LOCATION: 76644..76664	
31	OTHER INFORMATION: 20-853..pu	
32	NAME/KEY: primer_bind	
33	LOCATION: 77166..77185	
34	OTHER INFORMATION: 20-853..rp	complement
35	NAME/KEY: primer_bind	
36	LOCATION: 1220..1238	
37	OTHER INFORMATION: 20-828-311..mis	
38	NAME/KEY: primer_bind	
39	LOCATION: 1240..1258	
40	OTHER INFORMATION: 20-828-311..mis	complement
41	NAME/KEY: primer_bind	
42	LOCATION: 12328..12346	
43	OTHER INFORMATION: 17-42-319..mis	
44	NAME/KEY: primer_bind	
45	LOCATION: 12348..12366	
46	OTHER INFORMATION: 17-42-319..mis	complement
47	NAME/KEY: primer_bind	
48	LOCATION: 15222..15240	
49	OTHER INFORMATION: 17-41-250..mis	
50	NAME/KEY: primer_bind	
51	LOCATION: 15242..15260	
52	OTHER INFORMATION: 17-41-250..mis	complement
53	NAME/KEY: primer_bind	
54	LOCATION: 42199..42217	
55	OTHER INFORMATION: 20-841-149..mis	
56	NAME/KEY: primer_bind	
57	LOCATION: 42219..42237	
58	OTHER INFORMATION: 20-841-149..mis	complement
59	NAME/KEY: primer_bind	
60	LOCATION: 45423..45441	
61	OTHER INFORMATION: 20-842-115..mis	
62	NAME/KEY: primer_bind	
63	LOCATION: 45443..45461	
64	OTHER INFORMATION: 20-842-115..mis	complement
65	NAME/KEY: primer_bind	
66	LOCATION: 77039..77057	
67	OTHER INFORMATION: 20-853-415..mis	

```

NAME/KEY: primer_bind
LOCATION: 77059..77077
OTHER INFORMATION: 20-853-415. mis complement
NAME/KEY: misc_binding
LOCATION: 1227..1251
OTHER INFORMATION: 20-828-311. probe
NAME/KEY: misc_binding
LOCATION: 12355..12359
OTHER INFORMATION: 17-42-319. probe
NAME/KEY: misc_binding
LOCATION: 15229..15253
OTHER INFORMATION: 17-41-250. probe
NAME/KEY: misc_binding
LOCATION: 42206..42230
OTHER INFORMATION: 20-841-149. probe
NAME/KEY: misc_binding
LOCATION: 45430..45454
OTHER INFORMATION: 20-842-115. probe
NAME/KEY: misc_binding
LOCATION: 77046..77070
OTHER INFORMATION: 20-853-415. probe
US-09-751-877-1

```

Alignment Scores:		
Pred. No.:	9,21e-296	81001
Score:	312.00	312
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	85.25%	Indels: 0
DB:	10	Gaps: 0

QY	55	THLEALYSASPSERLEUUGINMSPLEUASNAEMEAASLYSHEUUGIULYSLEU	74
Db	14272	ACCCTGMAAGACACGCTTAGACGAAAGCCTCAACATATATACAAAGTCTCGAANAAGCTG	14331
QY	75	ATGProLeuSerGlySerGluAlaProArgLeuProGlnAspProValGlyMetArgArg	94
Db	14332	ACGCCTCTGAGTCGGACGACGAGGCTCTCCGCTCCACACAGACCCGGTGGCATGCGCGCG	14391
QY	95	GlnLeuGlnGlnGluLeuGlnGluValLysAlaArgLeuGlnProTyrMetAlaGluAla	114
Db	14392	CAGCTGCAGACGAGAGTGTGAGGAGGTGAAGGCTCGCTCCACCCCTACATGGCAGAGGG	14451
QY	115	HisGluLeuValGlyTyrPheSerLeuGluGlyLeuArgGlnGlnLeuLysProTyrPheMet	134
Db	14452	CACGAGCTGTGGCTGGAAATTTGGAGGGCTTGCGGCACCACTGAAGGCCCTACACGATG	14511
QY	135	AspLeuMetGluGlnValAlaLeuValArgValGlnGluLeuGlnGluLeuValArgValVal	154
Db	14512	GATCTGATGACGACAGGCTGGCCCTCCGCGCTGCAGACGCTCCAGACGACATGGCCGCGTG	14571
QY	155	GlyGlyAspThrLysAlaGlnLeuLeuGlnGlyGlyValAspGluAlaTrpAlaLeuLeuGln	174
Db	14572	GGGGAAGACACCAAGGCCACAGTTGCTGGGGGGCGGTGGAGAGGCTTGCGCTTGTGCGAG	14631
QY	175	GlyLeuGlnSerArgValValHisHisThrGlyValArgPheLysGluLeuPheHisProTyr	194
Db	14632	GGAGCTGCACAGCGCGCTGGTGGACCAACACGCGCGCTTCAAAAGACTCTTCAACCCATAC	14691
QY	195	AlaGluSerLeuValSerGlyIleGlyArgHisValGlnGluLeuHisArgSerValAla	214
Db	14692	GCGGAGAGCGCTGTGAGCGGCATCGGCGGCCACGTCGACGAGACTGCACCGCAGCTGGCT	14751
QY	215	ProHisAlaProAlaSerProAlaArgLeuSerArgCysValGlnValLeuSerArgLys	234
Db	14752	CCGCAAGCCCCCGCGACGCCCGCGCGCTCACTGCTCGTCCAGGTGCTCTCCCGGAAG	14811
QY	235	LeuThrLeuLysAlaLysAlaLeuHisAlaArgIleGlnGlnHisLeuAspGlnLeuArg	254
Db	14812	CTTCAGGCTCAAGGCCCAAGGCGCTGCACGACGCATCTCACACAAACCTGGACAGCTGGCC	14871


```

DB 124 CACGAGAGCTTGAGAGAGCGTGAAGCTCGGCTTCCACCCCTACATGGAGAGCCGACAGAC
QY 117 LeuValGlyTrpAsnLeuGlnGlyLeuArgGlnGlnLeuysProGlyThrMetAspLeu
DB 184 CTGGTGGCTGCAATTGGAGAGCTTCCGCA-CAACTGAACCCCTACACATGATGATCTG
QY 137 MetGlnGlnValAlaLeuArgValGlnGlnLeuGlnGlnLeuArgValGlnGlnGln
DB 243 ATGGACAGAGTGGCCCTGGCGTGCAGAGCTGACAGAGAGTGGCGCTGGGGGAA
QY 157 AspThrIysAlaGlnLeuLeuGlnGlyValAspGluAlaIrrpAlaLeuGlnGlnGln
DB 303 GACACCAAGGCCCGCTTGGCTGGCGCCCTGGAGAGAGCTTGGCTTGGCTGCAAGGAGCTG
QY 177 GlnSerArgValValHisIsthrGlyArgPheGlyLeuPhe
DB 363 CACAGCGCGCTGCTGCACACACCGCCGCTTCAAGACCTCTTC 407

```

RESULT 9

```

US-09-800-729-210
; Sequence 210, Application US/09800729
; Patent No. US20020068319A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 210
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-800-729-210

```

```

Alignment Scores:
Pred. No.: 6,97e-24 Length: 331
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.02% Indels: 0
DB: 10 Gaps: 0

```

US-09-842-364-3 (1-366) x US-09-800-729-210 (1-331)

```

QY 334 GLLYValLeuSerLysLeuGlnAlaArgLeuAspLeuTrpGluAspIleThrHis 353
DB 1 CGCAAGCTTCTGCAACCTGCAAGCCCGCTCTGATGACCTTGGGACACATCACTAC 60

```

```

QY 354 SerLeuHisAspGlnGlnHisSerHisLeuGlyAspPro 366
DB 61 AGCCTCATGACCAAGGCGCACAGCATCTGGGGGACCCC 99

```

RESULT 10

```

US-09-983-965-4674
; Sequence 4674, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26

```

```

; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 4674
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 14-LIB34-048-Q1-EI-D5
US-09-983-965-4674

```

US-09-842-364-3 (1-366) x US-09-983-965-4674 (1-301)

```

QY 8 LeuThrTrpAlaLeuAlaLeuSer 16
DB 31 CTGACCTGGCTCTGCGCCCTCTCTCA 57

```

RESULT 11

```

US-09-983-965-4360
; Sequence 4360, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 4360
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 02-LIB34-007-Q1-EI-A9
US-09-983-965-4360

```

Alignment Scores:

```

Pred. No.: 3.11 Length: 397
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 10 Gaps: 0

```

US-09-842-364-3 (1-366) x US-09-983-965-4360 (1-397)

```

QY 8 LeuThrTrpAlaLeuAlaLeuSer 16
DB 25 CTGACCTGGCTCTGCGCCCTCTCTCA 51

```

RESULT 12

```

US-09-960-352-10825
; Sequence 10825, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:

```

APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 10825
LENGTH: 405
TYPE: DNA
ORGANISM: Bos laurus
OTHER INFORMATION: Clone ID: 46-11B34-061-Q1-E1-D6
US-09-960-352-10825

Alignment Scores:
Pred. No.: 3.16 Length: 405
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: Gaps: 0

US-09-842-364-3 (1-366) x US-09-960-352-10825 (1-405)

OY 95 GlnLeuGlnGluGluLeuGluGluVal 103
DB 241 CAGCTGCAGAGAGAGCTGAGAGAGTG 267

RESULT 13
US-09-960-352-8211
Sequence 8211, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 8211
LENGTH: 419
TYPE: DNA
ORGANISM: Bos laurus
FEATURE:
NAME/KEY: unsure
LOCATION: (386)
OTHER INFORMATION: unsure at all n locations
US-09-960-352-8211

Alignment Scores:
Pred. No.: 3.26 Length: 419
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: Gaps: 0

US-09-842-364-3 (1-366) x US-09-960-352-8211 (1-419)

OY 8 LeuThrTrpAlaLeuAlaLeuLeuSer 16
DB 31 CTGACCTGGGCTCTGCCCTCTCTCA 57

RESULT 14
US-09-938-842A-2586

Sequence 2586, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krebs, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SAME, AND METHODS OF USE
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2586
LENGTH: 822
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2586

Alignment Scores:
Pred. No.: 6.06 Length: 822
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: Gaps: 0

US-09-842-364-3 (1-366) x US-09-938-842A-2586 (1-822)

OY 95 GlnLeuGlnGluGluLeuGluGluVal 103
DB 533 CAGCTGCAGAGAGAGCTGAGAGAGATT 559

RESULT 15
US-09-815-242-7404/C
Sequence 7404, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7404
 LENGTH: 2850
 TYPE: DNA
 ORGANISM: Helicobacter pylori
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(2850)
 US-09-815-242-7404

Alignment Scores:
 Pred. No.: 19
 Score: 9.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 2.46%
 DB: 10
 Length: 2850
 Matches: 9
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-842-364-3 (1-366) x US-09-815-242-7404 (1-2850)

Qy 171 AAlaLeuLeuGlnGlyLeuGlnSerArq 179
 |||||
 Db 1135 GCACCTCTTCACGCGATTGCCATCCGCC 1109

Search completed: January 23, 2003, 19:25:28
 Job time : 89 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2003, 16:10:07 : Search time 62 Seconds
(without alignments)
786.608 Million cell updates/sec

Title: US-09-842-364-3

Perfect score: 1877

Sequence: 1 MASMAVLTWALTLALSAFSA.....LMEDITRSHDGHSHLGD 366

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: A_Geneseq_101002: *
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
24: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1877	100.0	366	21	AA825778 Human secreted pro
2	1877	100.0	366	22	AAE11932 Human C6122 (or C8
3	1877	100.0	366	22	AAU14554 Human novel protei
4	1877	100.0	366	22	AA820103 Human apolipoprote
5	1877	100.0	366	22	AA875364 Human secreted pro
6	1877	100.0	400	22	AAE11941 Human lipid metabo
7	1877	100.0	400	22	AAU14318 Human novel protei
8	1871	99.7	366	22	AA890666 Human secreted pro
9	1864	99.3	363	21	AAV54552 Amino acid sequenc
10	1858	99.0	363	22	AA890573 Human secreted pro

11	1858	99.0	363	23	AB65475 Human albumin fus1
12	1396	74.4	274	23	AAU86153 Human PRO773 polyP
13	1381	73.6	274	23	ABG34037 Human pro peptidase
14	1293	68.9	367	21	AAV54561 Amino acid sequenc
15	1293	68.9	367	21	AAV67399 Rat rap3 protein s
16	350	18.6	382	22	AA890663 Human secreted pro
17	328	17.5	396	22	AA890664 Human secreted pro
18	324	17.3	396	23	AAU10864 Human apolipoprote
19	322.5	17.2	396	23	AAU10867 Human apolipoprote
20	321	17.1	396	23	AAU10860 Human apolipoprote
21	321	17.1	396	23	AAU10865 Human apolipoprote
22	321	17.1	396	23	AAU10866 Human apolipoprote
23	321	17.1	396	23	AAU10868 Human apolipoprote
24	321	17.1	396	23	AAU10869 Human apolipoprote
25	320	17.0	396	23	AAU10861 Human apolipoprote
26	320	17.0	396	23	AAU10861 Human apolipoprote
27	318	16.9	396	23	AAU10861 Human apolipoprote
28	317	16.9	396	23	AAU10862 Human apolipoprote
29	316	16.8	396	23	AAU10862 Human apolipoprote
30	316	16.8	396	23	AAU10862 Human apolipoprote
31	316	16.8	396	23	AAU10862 Human apolipoprote
32	316	16.8	396	23	AAU10862 Human apolipoprote
33	315	16.8	396	23	AAU10862 Human apolipoprote
34	315	16.8	396	23	AAU10862 Human apolipoprote
35	313	16.7	396	23	AAU10862 Human apolipoprote
36	310	16.5	396	23	AAU10862 Human apolipoprote
37	309	16.5	396	23	AAU10862 Human apolipoprote
38	309	16.5	396	23	AAU10862 Human apolipoprote
39	308.5	16.4	396	23	AAU10862 Human apolipoprote
40	308.5	16.4	396	23	AAU10862 Human apolipoprote
41	300	16.0	396	23	AAU10862 Human apolipoprote
42	300	16.0	396	23	AAU10862 Human apolipoprote
43	298.5	15.9	391	22	AA890665 Human apolipoprote
44	298	15.9	391	22	AA890665 Human apolipoprote
45	292	15.6	393	14	AA894940 Human apolipoprote

ALIGNMENTS

RESULT 1	AA825778	standard; Protein: 366 AA.
XX	AA825778:	
AC	AA825778:	
AC	AA825778:	
DT	28-NOV-2000	(first entry)
XX	Human secreted protein SEQ ID #90.	
XX	Human: secreted protein; forensic procedure; gene therapy;	
KW	chromosome mapping; cancer; autoimmune disease; cardiovascular disorder;	
KW	cystic fibrosis; hypothyroidism; immunological disorder; amyloidosis;	
KW	brain disorder; skeletal muscle disorder; eye disorder; obesity;	
KW	mitochondriopathy; diabetes; atherosclerosis; Alzheimer's disease;	
KW	neurodegenerative disorder; graft rejection; dementia; hyperlipidemia;	
KW	septic shock; impotence.	
OS	Homo sapiens.	
XX	MO200037491-A2.	
XX	29-JUN-2000.	
XX	20-DEC-1999;	99WO-1802058.
XX	22-DEC-1998;	98US-0113686.
XX	25-JUN-1999;	99US-0141032.
XX	(GEST) GENSET.	
XX	Bouguetier L, Dumas J, Duclert A;	
XX		

Query Match	100.0%	Score 1877	DB 21	Length 366
Best Local Similarity	100.0%	Pred. No. 3.9e-155		
Matches 366	Conservative 0	Mismatches 0	Indels 0	Gaps 0
1 MASMAAVLTWALALLSASFATQARKGFMDYFQTSQDKGRVQDIHOOKMARREPARTIKDSL	1	1	1	1
1 MASMAAVLTWALALLSASFATQARKGFMDYFQTSQDKGRVQDIHOOKMARREPARTIKDSL	1	1	1	1
61 EODLNNMNFLEKLRPLSGSEAPRLRPQDVGMRKROLQELIEVKARLQPYMAEAHELQVM	61	61	61	61
61 EODLNNMNFLEKLRPLSGSEAPRLRPQDVGMRKROLQELIEVKARLQPYMAEAHELQVM	61	61	61	61
121 NLEGLRQOLKPYTMDLMEQVALRVQELQRLVNGEDTKAQLLGGVDEAMALLQGLQSLRV	121	121	121	121
121 NLEGLRQOLKPYTMDLMEQVALRVQELQRLVNGEDTKAQLLGGVDEAMALLQGLQSLRV	121	121	121	121
181 VHHTRGFKLFFPYAESLVSIGLRHVOELHRSVAPHPAPSPARLSKCVQVLSRKLTLMKAK	181	181	181	181
181 VHHTRGFKLFFPYAESLVSIGLRHVOELHRSVAPHPAPSPARLSKCVQVLSRKLTLMKAK	181	181	181	181
241 ALHAAIQONLDLRELSNAFAGTGTSEEGAGDPQMLSEVQRQLQAFQDITYLQIAAFT	241	241	241	241
241 ALHAAIQONLDLRELSNAFAGTGTSEEGAGDPQMLSEVQRQLQAFQDITYLQIAAFT	241	241	241	241
301 RAIDQETEEVQOQLAPPPGHSAPAFPEFQOTSGKVLSTLQARLDLMDIEDITHSLHDQGH	301	301	301	301
301 RAIDQETEEVQOQLAPPPGHSAPAFPEFQOTSGKVLSTLQARLDLMDIEDITHSLHDQGH	301	301	301	301
361 SHLQDP 366	361	361	361	361
361 SHLQDP 366	361	361	361	361

XX	Human: apolipoprotein: lipase: lipoprotein receptor; ALR: angina;
KM	cardiovascular disease; lipid metabolism; myocardial infarction;
KM	cerebral ischemia; arterial thrombosis; thrombolytic; atherosclerotic;
KM	coronary artery thrombosis; cerebral artery thrombosis; stroke;
KM	intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
XX	neuroprotectant; cerebroprotective.
OS	Homo sapiens.
XX	
EH	Key
FT	Location/Qualifiers
FT	Peptide
FT	1..23
FT	/label= Signal_peptide
FT	Protein
FT	24..366
FT	/note= "Human mature CG122 (or C868) protein"
FT	Domain
FT	4..251
FT	/note= "Apolipoprotein A1/A4/E family domain"
FT	Domain
FT	75..130
FT	/note= "Apolipoprotein plasma lipid transport domain"
FT	92..142
FT	/note= "Apolipoprotein E precursor domain"
PN	Domain
XX	MO200179446-A2.
PD	
XX	25-OCT-2001.
PF	
XX	16-APR-2001; 2001MO-US12529.
PR	
XX	14-APR-2000; 2000US-197137P.
PR	20-JUN-2000; 2000US-0596042.
PR	03-AUG-2000; 2000US-0631451.
PR	22-SEP-2000; 2000US-0667298.
PR	17-NOV-2000; 2000US-0714936.
XX	
PA	(HYSE-) HYSEQ INC.
PI	
PI	Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI	Liu C, Asundi V, Zhao QA, Wehrman T, Dirmannac RT, Ken F, Qian XB;
PI	Wang D;
DR	
DR	WPI: 2001-611724/70.
DR	N-PSDB: AAD19216.
XX	
PT	Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT	receptor polypeptides, useful for preventing diagnosing and treating
PT	lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX	
PS	Claim 10; Page 148-150; 266pp; English.
XX	
CC	The invention relates to polynucleotides encoding proteins CG122, CG129
CC	CG95, CG121, CG152, CG27, CG153 and CG168 which are related to proteins
CC	involved in lipid metabolism and cardiovascular disease such as human
CC	apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC	and protein sequences are useful for treating or preventing disorders
CC	associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
CC	expression and for treating lipid metabolism, cardiovascular diseases
CC	and thrombosis. Antibodies against these proteins are useful for
CC	determining the presence of or predisposition to a disease associated
CC	with altered levels of these sequences. ALR polypeptides are also
CC	useful for identifying agents (agonists and antagonists) that bind to
CC	them and cells expressing ALR proteins are useful for identifying a
CC	therapeutic agent for use in treatment of a pathology related to
CC	aberrant expression or physiological interactions of this polypeptide.
CC	Vectors comprising these DNA and protein sequences are also useful for
CC	producing ALR proteins. The nucleic acids and polypeptides of the
CC	invention are also useful for the treatment of occlusive cardiovascular
CC	diseases, myocardial infarction, cerebral ischemia, angina, arterial
CC	thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC	or intracardiac thrombosis and stroke. The nucleotides of the invention
CC	are used in gene therapy. The present sequence is human CG122 (or C868)
CC	protein.
XX	
SO	Sequence 366 AA:

Human: apolipoprotein; lipase; lipoprotein receptor; ALR; angina; cardiovascular disease; lipid metabolism; myocardial infarction; cerebral ischemia; arterial thrombosis; thrombolytic; antithrombotic; coronary artery thrombosis; cerebral artery thrombosis; stroke; intracardiac thrombosis; gene therapy; cardiovascular; vasodilator; neuroprotectant; cerebroprotective.

Homo sapiens.

Key	Location/Qualifiers
Peptide	1..23
Protein	/label= Signal _peptide 24..366
Domain	/note= "Human mature CG122 (or C86) protein"
Domain	4..251
Domain	/note= "Apolipoprotein AI/A4/E family domain"
Domain	75..130
Domain	/note= "Apolipoprotein plasma lipid transport domain"
Domain	92..142
Domain	/note= "Apolipoprotein E precursor domain"

MO200179446-A2.

25-OCT-2001.

16-APR-2001; 2001WO-US12529.

14-APR-2000; 2000US-197137P.
20-JUN-2000; 2000US-0598042.
03-AUG-2000; 2000US-0631451.
22-SEP-2000; 2000US-0667298.
17-NOV-2000; 2000US-0714936.

(HYSE-) HYSEQ INC.

Ballingier DG, Joeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R; Liu C, Asundi V, Zhao QA, Wehrman T, Dymnac RT, Ken F, Qian XB; Wang D:

WPI: 2001-611724/70.
N-PSDB: AAD19216.

Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein receptor polypeptides, useful for preventing diagnosing and treating lipid metabolism disorders, thrombosis and cardiovascular diseases -

Claim 10; Page 148-150; 266pp; English.

The invention relates to polynucleotides encoding proteins CG122, CG179 CG95, CG121, CG162, CG27, CG153 and CG166 which are related to proteins involved in lipid metabolism and cardiovascular disease such as human apolipoproteins, lipases and lipoprotein receptor proteins. These DNA and protein sequences are useful for treating or preventing disorders associated with apolipoproteins, lipases and lipoprotein receptor (ALR) expression and for treating lipid metabolism, cardiovascular diseases and thrombosis. Antibodies against these proteins are useful for determining the presence of or predisposition to a disease associated with altered levels of these sequences. ALR polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them and cells expressing ALR proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of this polypeptide. Vectors comprising these DNA and protein sequences are also useful for producing ALR proteins. The nucleic acids and polypeptides of the invention are also useful for the treatment of occlusive cardiovascular diseases, myocardial infarction, cerebral ischemia, angina, arterial thrombosis, coronary artery thrombosis and cerebral artery thrombosis or intracardiac thrombosis and stroke. The nucleotides of the invention are used in gene therapy. The present sequence is human CG122 (or C86) protein.

sequence 366 AA:

Query Match 100.0%: Score 1877; DB 22: Length 366;
 Best Local Similarity 100.0%: Pred. No. 3.9e-155;
 Matches 366: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASMAAVLTWALALLSFASTQARKGFWDFYSGTSGDKGVEQIHQOKMAREPATLKDSL 60
 |||||||
 DB 1 MASMAAVLTWALALLSFASTQARKGFWDFYSGTSGDKGVEQIHQOKMAREPATLKDSL 60
 |||||||

QY 61 EODLNMMNKFLEKLRPLSSSEAPRLPQDPVGMRRQLOELEEYKARLQPYMAEAHELVCW 120
 |||||||
 DB 61 EODLNMMNKFLEKLRPLSSSEAPRLPQDPVGMRRQLOELEEYKARLQPYMAEAHELVCW 120
 |||||||

QY 121 NIEGLRQOLKPYTMDLMEQVALRVQELQVGEDETKAQLLGVDENALLQGLQSRV 180
 |||||||
 DB 121 NIEGLRQOLKPYTMDLMEQVALRVQELQVGEDETKAQLLGVDENALLQGLQSRV 180
 |||||||

QY 181 VHHGTGRKELFHPYAESLVSIGIRHVOELHRSVAPIAPASPARLSRCVOYLSKRLTLKAK 240
 |||||||
 DB 181 VHHGTGRKELFHPYAESLVSIGIRHVOELHRSVAPIAPASPARLSRCVOYLSKRLTLKAK 240
 |||||||

QY 241 ALHARIQONLDOLRELSRAFACTGTEGAGPRPOMLSEEVROQLAFRODTYLOIAAFT 300
 |||||||
 DB 241 ALHARIQONLDOLRELSRAFACTGTEGAGPRPOMLSEEVROQLAFRODTYLOIAAFT 300
 |||||||

QY 301 RAIDQETEEVQOOLAPPPGHSAPAFEFQOTDSGKVLSKLOARLDLMDLMDITSHLDQGH 360
 |||||||
 DB 301 RAIDQETEEVQOOLAPPPGHSAPAFEFQOTDSGKVLSKLOARLDLMDLMDITSHLDQGH 360
 |||||||

QY 361 SHLGDP 366
 |||||||
 DB 361 SHLGDP 366
 |||||||

RESULT 3
 AAU14554
 ID AAU14554 standard; Protein: 366 AA.

XX Homo sapiens.
 PN WO20015437-A2.
 XX 02-AUG-2001.
 PD 25-JAN-2001: 2001WO-US02623.
 XX 25-JAN-2001: 2000US-0491404.
 PR (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI WPI: 2001-451939/48.
 DR N-PSDB: AAS22859.
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage -
 XX Example 4: Page 871-872; 894pp; English.

XX The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour. In assays to determine biological activity, to
 CC raise antibodies/elicite an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis.
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.

Sequence 366 AA:
 SQ

Query Match 100.0%: Score 1877; DB 22: Length 366;
 Best Local Similarity 100.0%: Pred. No. 3.9e-155;
 Matches 366: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASMAAVLTWALALLSFASTQARKGFWDFYSGTSGDKGVEQIHQOKMAREPATLKDSL 60
 |||||||
 DB 1 MASMAAVLTWALALLSFASTQARKGFWDFYSGTSGDKGVEQIHQOKMAREPATLKDSL 60
 |||||||

QY 61 EODLNMMNKFLEKLRPLSSSEAPRLPQDPVGMRRQLOELEEYKARLQPYMAEAHELVCW 120
 |||||||
 DB 61 EODLNMMNKFLEKLRPLSSSEAPRLPQDPVGMRRQLOELEEYKARLQPYMAEAHELVCW 120
 |||||||

QY 121 NIEGLRQOLKPYTMDLMEQVALRVQELQVGEDETKAQLLGVDENALLQGLQSRV 180
 |||||||
 DB 121 NIEGLRQOLKPYTMDLMEQVALRVQELQVGEDETKAQLLGVDENALLQGLQSRV 180
 |||||||

QY 181 VHHGTGRKELFHPYAESLVSIGIRHVOELHRSVAPIAPASPARLSRCVOYLSKRLTLKAK 240
 |||||||
 DB 181 VHHGTGRKELFHPYAESLVSIGIRHVOELHRSVAPIAPASPARLSRCVOYLSKRLTLKAK 240
 |||||||

QY 241 ALHARIQONLDOLRELSRAFACTGTEGAGPRPOMLSEEVROQLAFRODTYLOIAAFT 300
 |||||||
 DB 241 ALHARIQONLDOLRELSRAFACTGTEGAGPRPOMLSEEVROQLAFRODTYLOIAAFT 300
 |||||||

QY 301 RAIDQETEEVQOOLAPPPGHSAPAFEFQOTDSGKVLSKLOARLDLMDLMDITSHLDQGH 360
 |||||||
 DB 301 RAIDQETEEVQOOLAPPPGHSAPAFEFQOTDSGKVLSKLOARLDLMDLMDITSHLDQGH 360
 |||||||

QY 361 SHLGDP 366
 |||||||
 DB 361 SHLGDP 366
 |||||||

RESULT 4
 AAB20103
 ID AAB20103 standard; Protein: 366 AA.

XX AAB20103;
 XX 23-APR-2001 (first entry)
 DT Human apolipoprotein A-IV-related protein (A44RP).
 XX Human apolipoprotein A-IV-related protein; A44RP; human;
 DE Apolipoprotein A-IV-related protein; A44RP; human;
 XX

bi-allelic marker: lipid metabolism; liver related disorder;
 obesity; diabetes; coronary heart disease; diagnosis: therapy.

Homo sapiens.

MO200100803-A2.

04-JAN-2001.

21-JUN-2000; 2000MO-1B01011.

25-JUN-1999; 99US-0141032.

20-DEC-1999; 99MO-1B02058.

21-DEC-1999; 99US-0469099.

(GEST) GENSET.

Yen F, Denison B, Bour B, Bihain B, Bougueleret L, Duclert A;

Dumas Milne Edwards J;

WPI: 2001-071485/08

N-PSDB; AAF30035, AAF30036.

Claim 12; Page 240; 260pp; English.

The present sequence is that of the human apolipoprotein A-IV-related protein (AA4RP). AA4RP is differentially expressed in obese mouse models, indicating a role in lipid metabolism related disorders. It appears to be the human homologue of rat regeneration associated protein (RAP3), which is believed to be involved in liver regeneration, and is likely to have a similar function. It also shows 52% similarity to apolipoprotein A-IV, and is likely to have a similar function. The invention also provides AA4RP cDNAs (see AAF30036) and genomic DNA (see AAF30035), bi-allelic markers identified in the AA4RP gene and from genomic regions flanking the gene, and methods for genotyping a nucleic acid containing 1 or more of the bi-allelic markers. Also provided is a method for detecting a statistical correlation between a bi-allelic marker allele and a phenotype and/or between a bi-allelic marker haplotype and a phenotype. Diagnostic methods are provided for determining whether an individual is at risk of developing a lipid metabolism related disorder and/or a liver related disorder, such as obesity, diabetes and coronary heart disease, and whether the individual suffers from such a disorder as a result of a polymorphism in the AA4RP gene. AA4RP polypeptides can also be used in therapy, to screen for agonist and antagonist compounds, and to raise AA4RP-specific antibodies.

Sequence 366 AA;

Query Match 100.0%; Score 1877; DB 22; Length 366;
 Best Local Similarity 100.0%; Pred. No. 3.9e-155;
 Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MASMAAVLTWALALSAFATGARKGFMDFQSOTSGDKRGVQIHQOKMAREPATLKDSL 60
 1 MASMAAVLTWALALSAFATGARKGFMDFQSOTSGDKRGVQIHQOKMAREPATLKDSL 60
 61 EODLNMMNKFLEKLRPLSGSEAPRLPODPVGNRRQLQOELEEVKARLOPYMAEAHELIVGM 120
 61 EODLNMMNKFLEKLRPLSGSEAPRLPODPVGNRRQLQOELEEVKARLOPYMAEAHELIVGM 120
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 181 VHTGREFELFHPYAESLVSGIRHVOELHRSVAPHAPASPARLSRCVOVLSRKLTLLAK 240
 181 VHTGREFELFHPYAESLVSGIRHVOELHRSVAPHAPASPARLSRCVOVLSRKLTLLAK 240

241 ALHARIQOQLDOLRELSRAFACTGTEGAGPPQMLSEEVNQRLOAFRODTYLOIAFT 300
 241 ALHARIQOQLDOLRELSRAFACTGTEGAGPPQMLSEEVNQRLOAFRODTYLOIAFT 300
 301 RAIDETEVEVOQLAPPFGHSAPAFEPFOOTDSGKVLSTQARLDLMDITHSILIDGSH 360
 301 RAIDETEVEVOQLAPPFGHSAPAFEPFOOTDSGKVLSTQARLDLMDITHSILIDGSH 360
 361 SHLGD 366
 361 SHLGD 366

RESULT 5
 AAB75364
 ID AAB75364 standard; protein; 366 AA.

AC AAB75364;

DT 05-APR-2001 (first entry)

DE Human secreted protein #23.

KW Secreted protein; prevention; treatment; diagnosis; disease;

infection.

OS Homo sapiens.

PN WO200100806-A2.

PD 04-JAN-2001.

PF 21-JUN-2000; 2000MO-1B00951.

PR 25-JUN-1999; 99US-0141032.

PR 21-DEC-1999; 99US-0469099.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Bougueleret L, Jobert S;

WPI: 2001-071487/08.

PS Claim 10; Page 288-289; 307pp; English.

The present invention relates to 49 secreted proteins and the cDNAs encoding them. The protein and nucleic acids may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate protein expression.

Sequence 366 AA;

Query Match 100.0%; Score 1877; DB 22; Length 366;
 Best Local Similarity 100.0%; Pred. No. 3.9e-155;
 Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MASMAAVLTWALALSAFATGARKGFMDFQSOTSGDKRGVQIHQOKMAREPATLKDSL 60
 1 MASMAAVLTWALALSAFATGARKGFMDFQSOTSGDKRGVQIHQOKMAREPATLKDSL 60
 61 EODLNMMNKFLEKLRPLSGSEAPRLPODPVGNRRQLQOELEEVKARLOPYMAEAHELIVGM 120
 61 EODLNMMNKFLEKLRPLSGSEAPRLPODPVGNRRQLQOELEEVKARLOPYMAEAHELIVGM 120
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 121 NLEGLRQOLKPYTMDLMEQVALRVOELQRLRVGVEDTKAQLLGVDAMALLQGLQSRV 180
 181 VHTGREFELFHPYAESLVSGIRHVOELHRSVAPHAPASPARLSRCVOVLSRKLTLLAK 240
 181 VHTGREFELFHPYAESLVSGIRHVOELHRSVAPHAPASPARLSRCVOVLSRKLTLLAK 240

DB 181 VHHGTFEKELEPHVAESLVSGIRHVOELHRSVAPHAPASPARLSRCQVLSRKLTAK 240
 QY 241 ALHARIQONLDOLRELSRAFGCTGTCAGAPDPOMLSEEVQRORLOARODTYLQIAAFT 300
 DB 241 ALHARIQONLDOLRELSRAFGCTGTCAGAPDPOMLSEEVQRORLOARODTYLQIAAFT 300
 QY 301 RAIDQETEEVVOQLAPPPGHSAPAFEPFOQTDSGKVLKLAARDLDMEDITHSLHDGCH 360
 DB 301 RAIDQETEEVVOQLAPPPGHSAPAFEPFOQTDSGKVLKLAARDLDMEDITHSLHDGCH 360
 QY 361 SHLGD 366
 DB 361 SHLGD 366
 RESULT 6
 ID AE11941 standard; Protein: 400 AA.
 AC AE11941:
 DE 18-DEC-2001 (first entry)
 XX Human lipid metabolism related protein #4.
 XX Human: apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
 KM cardiovascular disease; lipid metabolism; myocardial infarction;
 KM cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KM coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KM intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KM neuroprotectant; cerebroprotective.
 XX Homo sapiens.
 OS
 PN MO200179446-A2.
 XX 25-OCT-2001.
 PD 16-APR-2001; 2001WO-US12529.
 PF 14-APR-2000; 2000US-197137P.
 PR 20-JUN-2000; 2000US-0598042.
 PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 XX (HXSE-) HXSEQ INC.
 PA Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 DR WPI: 2001-611724/70.
 DR N-PSDB: AAD19236.
 XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX Claim 10; Page 260-261; 266pp: English.
 PS The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to

CC them and cells expressing ALR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALR proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human protein
 CC related to proteins involved in lipid metabolism.
 CC
 XX Sequence 400 AA:
 SQ
 Query Match 100.0%; Score 1877; DB 22; Length 400;
 Best local similarity 100.0%; Pred. No. 4.4e-155;
 Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASMAAVLTWALALLSAFSAATQARKGFMDYFSQTSQDGRVEQIHQKMAREPATLKDSL 60
 DB 35 MASMAAVLTWALALLSAFSAATQARKGFMDYFSQTSQDGRVEQIHQKMAHPATLKDSL 94
 QY 61 EODLNNMKNKFLKLRPLSGSEAPRLPDPVCKRROLQELLEVKARLOPYMAEAHELVGW 120
 DB 95 EODLNNMKNKFLKLRPLSGSEAPRLPDPVCKRROLQELLEVKARLOPYMAEAHELVGW 154
 QY 121 NIEGIRQOLKPYTMDLMOVALRVOELRVGEEDTRKAOLIGVDHAMAALQGLQSRV 180
 DB 155 NIEGIRQOLKPYTMDLMOVALRVOELRVGEEDTRKAOLIGVDHAMAALQGLQSRV 214
 QY 181 VHHGTFEKELEPHVAESLVSGIRHVOELHRSVAPHAPASPARLSRCQVLSRKLTAK 240
 DB 215 VHHGTFEKELEPHVAESLVSGIRHVOELHRSVAPHAPASPARLSRCQVLSRKLTAK 274
 QY 241 ALHARIQONLDOLRELSRAFGCTGTCAGAPDPOMLSEEVQRORLOARODTYLQIAAFT 300
 DB 275 ALHARIQONLDOLRELSRAFGCTGTCAGAPDPOMLSEEVQRORLOARODTYLQIAAFT 334
 QY 301 RAIDQETEEVVOQLAPPPGHSAPAFEPFOQTDSGKVLKLAARDLDMEDITHSLHDGCH 360
 DB 335 RAIDQETEEVVOQLAPPPGHSAPAFEPFOQTDSGKVLKLAARDLDMEDITHSLHDGCH 394
 QY 361 SHLGD 366
 DB 395 SHLGD 400
 RESULT 7
 ID AAU14318 standard; Protein: 400 AA.
 AC AAU14318:
 DE 24-OCT-2001 (first entry)
 XX Human novel protein #189.
 XX Human: novel protein; antianemic; osteopathic; antiinflammatory;
 KM immunomodulatory; cytosolic; neuroprotective; vulnary; nootrophic;
 KM anticonvulsant; antiaftritic; cerebroprotective; antitumor; antiviral;
 KM antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;
 KM thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KM Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KM tissue regeneration; immune disorder.
 XX Homo sapiens.
 OS
 PN WO200155437-A2.
 XX 02-AUG-2001.
 PD 25-JAN-2001; 2001WO-US02623.
 PF

PR 25-JAN-2000; 2000US-0491404.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Dmanac RT;
 XX WPI: 2001-451939/48.
 DR N-PSDB: AAS22623.
 XX
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 XX nervous system disorders, and for regenerating bone and cartilage -
 XX
 PS Example 4; Page 629-630; 894pp; English.
 CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumour, in assays to determine biological activity, to
 CC raise antibodies/elicitor an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.
 CC
 XX
 SQ Sequence 400 AA:
 Query Match 100.0%; Score 1877; DB 22; Length 400;
 Best local Similarity 100.0%; Pred. No. 4.4e-155;
 Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASMAAVLTWALALSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSL 60
 DB 35 MASMAAVLTWALALSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSL 94
 QY 61 EODLNMMNFKLEKRLPLSGSEAPRLPODPVGMROLOEELVKKARLOPYMAEAEHLVGM 120
 DB 95 EODLNMMNFKLEKRLPLSGSEAPRLPODPVGMROLOEELVKKARLOPYMAEAEHLVGM 154
 QY 121 NLEGLRQOLKPYTMDLMEQVALRVOELQVQEDTAKQLLGVDVEMALLQGLQSRV 180
 DB 155 NLEGLRQOLKPYTMDLMEQVALRVOELQVQEDTAKQLLGVDVEMALLQGLQSRV 214
 QY 181 VHTGSRKELFHPYAESIVSGTGRHOELHRSVAPAPASPARLSQVYLSKRLTKAK 240
 DB 215 VHTGSRKELFHPYAESIVSGTGRHOELHRSVAPAPASPARLSQVYLSKRLTKAK 274
 QY 241 ALHARIDONLDQRELSRAFGTTEGAGPDPOMLSEEVRORLQAFRODTYLAQIAAF 300
 DB 275 ALHARIDONLDQRELSRAFGTTEGAGPDPOMLSEEVRORLQAFRODTYLAQIAAF 334
 QY 301 RAIDQEEVEVQOOLAPPHPGHSAPAFBFOOTDSCGKVLSTKQALDLMDLWEDITHSLDQGH 360
 DB 335 RAIDQEEVEVQOOLAPPHPGHSAPAFBFOOTDSCGKVLSTKQALDLMDLWEDITHSLDQGH 394
 QY 361 SHLGGP 366
 DB 395 SHLGGP 400

RESULT 8
 AAB90666
 ID AAB90666 standard: Protein: 366 AA.
 XX
 AC AAB90666;
 XX
 DT 01-JUN-2001 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 212.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic;
 KW dermatological; anti-inflammatory; anti-HIV; cytostatic; cardiant;
 KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;
 KW neurotropic; anticonvulsant; anti-Alzheimers; antiparkinsonian;
 KW antimicrobial; vulnerable; vaccine; gene therapy; cancer;
 KW protein coordinate data; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200121658-A1.
 XX
 PD 29-MAR-2001.
 XX
 XX 22-SEP-2000; 2000WO-US26013.
 XX
 PR 24-SEP-1999; 99US-0155709.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsu S, CA;
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
 PI Young PE, Wei P, Florence KA;
 XX
 DR WPI: 2001-235311/24.
 XX
 PT Nucleic acids encoding 32 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 PS Disclosure; Fig 7; 890pp; English.
 XX
 CC The present sequence is provided in a specification relating to nucleic
 CC acid molecules encoding 32 novel human secreted polypeptides. The nucleic
 CC acid molecules and polypeptides may be used in the prevention, diagnosis
 CC and treatment of diseases such as immune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus and human immuno-deficiency virus
 CC (HIV) infections), hyperproliferative disorders (e.g. cancers and
 CC Gaucher's disease), cardiovascular diseases (e.g. Schmitz syndrome,
 CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic
 CC disorders (e.g. corneal graft neovascularisation and diabetic
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or
 CC for promoting wound healing, regeneration and/or chemotaxis. The nucleic
 CC acid molecules may be used to produce the secreted polypeptides. They may
 CC also be used as DNA probes in diagnostic assays to detect and quantitate
 CC the presence of similar nucleic acid sequences in samples. The
 CC polypeptides may be used as antigens in the production of antibodies and
 CC in assays to identify modulators of their expression and activity.
 CC
 XX
 SQ Sequence 366 AA:
 Query Match 99.7%; Score 1871; DB 22; Length 366;
 Best local Similarity 99.7%; Pred. No. 1.3e-154;
 Matches 365; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MASMAAVLTWALALSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSL 60
 DB 1 MASMAAVLTWALALSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSL 60
 QY 61 EODLNMMNFKLEKRLPLSGSEAPRLPODPVGMROLOEELVKKARLOPYMAEAEHLVGM 120

Db 61 EDDLNNNNKFLKRLPLSGSEAPRLPDDPVGMRRQLQEELEEVKARLQPYMAEAHELVCW 120
 QY 121 NIEGLRQOLKPYTMDLMEQVALRVQELQEDTKAQLLGGVDEAMALLQGLQSRV 180
 Db 121 NIEGLRQOLKPYTMDLMEQVALRVQELQEDTKAQLLGGVDEAMALLQGLQSRV 180
 QY 181 VHTGRKELFHPYAESLVSGIGHVOELHRSVAHPAPASPARLSRCVQLSRKLTLLKAK 240
 Db 181 VHTGRKELFHPYAESLVSGIGHVOELHRSVAHPAPASPARLSRCVQLSRKLTLLKAK 240
 QY 241 ALHARIQONLDOLREELSRAPAGTGTGEGAGPPOMLSEEVROQLQAFRODTYLOIAAF 300
 Db 241 ALHARIQONLDOLREELSRAPAGTGTGEGAGPPOMLSEEVROQLQAFRODTYLOIAAF 300
 QY 301 RAIDQETEEVQOOLAPPFGHSAFAPEFOOTDSGKVLKQLARLDLMDLITSLHDQGH 360
 Db 301 RAIDQETEEVQOOLAPPFGHSAFAPEFOOTDSGKVLKQLARLDLMDLITSLHDQGH 360
 QY 361 SHLQDP 366
 Db 361 SHLQDP 366
 RESULT 9
 AAY54562
 ID AAY54562 standard; protein: 363 AA.
 AC AAY54562;
 XX 25-APR-2000 (first entry)
 DT
 DE Amino acid sequence of the human RAP3 protein.
 XX RAP3; regeneration association protein 3; liver regeneration;
 KM liver proliferation.
 XX Homo sapiens.
 OS
 PN WO200003013-A2.
 XX 20-JAN-2000.
 PD
 XX 12-JUL-1999; 99WO-EP04938.
 PR 10-JUL-1998; 98EP-0202336.
 XX
 PA (AMST-) AMSTERDAM MOLECULAR THERAPEUTICS BV.
 PI Chamuleau RAFM, Groenink M, Van Der Vijet HN, Iecowater ACJ;
 XX WPI: 2000-147615/13.
 DR N-PSDB: AA245874, AA245875.
 XX
 PT Isolated RAP3 gene, protein and antibody useful for diagnosing liver
 PS regeneration and/or cell proliferation -
 XX
 XX Claim 6; Fig 7; 42pp; English.
 CC The present sequence represents a human RAP3 (regeneration association
 CC protein 3). The RAP3 cDNA sequence was isolated from a human liver cDNA
 CC library. The RAP3 gene is involved in regeneration processes of the
 CC liver. The rat RAP3 gene was found to be upregulated 6 hours after
 CC partial hepatectomy, after which it was downregulated. The RAP3 cDNA
 CC sequence is useful as a source for PCR primers and probes, which are
 CC useful for detecting nucleotide sequences in a source material. The
 CC RAP3 cDNA sequence is useful as a marker of liver proliferation. The
 CC RAP3 protein is useful for the diagnosis of liver regeneration and liver
 CC cell proliferation. RAP3 antibodies, PCR primers and probes are useful
 CC for detecting the occurrence of liver cell proliferation in a patient.
 CC The RAP3 protein is also useful for enhancing the growth of
 CC regeneration of liver tissue comprising treating the liver tissue
 CC such as extracorporeal or intracorporeal.

SQ Sequence 363 AA;
 Query Match 99.3%; Score 1864; DB 21; Length 363;
 Best Local Similarity 100.0%; Pred. No. 5.3e-154;
 Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 MAAVLTWALALLSFASTQARKGFMDYFSQTSQDKGVEDIHOOKMAREPATLKDSLEOD 63
 Db 1 MAAVLTWALALLSFASTQARKGFMDYFSQTSQDKGVEDIHOOKMAREPATLKDSLEOD 60
 QY 64 LNNNNKFLKRLPLSGSEAPRLPDDPVGMRRQLQEELEEVKARLQPYMAEAHELVCW 123
 Db 61 LNNNNKFLKRLPLSGSEAPRLPDDPVGMRRQLQEELEEVKARLQPYMAEAHELVCW 120
 QY 124 GLRQOLKPYTMDLMEQVALRVQELQEDTKAQLLGGVDEAMALLQGLQSRV 183
 Db 121 GLRQOLKPYTMDLMEQVALRVQELQEDTKAQLLGGVDEAMALLQGLQSRV 180
 QY 184 TGRKELFHPYAESLVSGIGHVOELHRSVAHPAPASPARLSRCVQLSRKLTLLKAK 243
 Db 181 TGRKELFHPYAESLVSGIGHVOELHRSVAHPAPASPARLSRCVQLSRKLTLLKAK 240
 QY 244 ARIQONLDOLREELSRAPAGTGTGEGAGPPOMLSEEVROQLQAFRODTYLOIAAF 303
 Db 241 ARIQONLDOLREELSRAPAGTGTGEGAGPPOMLSEEVROQLQAFRODTYLOIAAF 300
 QY 304 DQETEEVQOOLAPPFGHSAFAPEFOOTDSGKVLKQLARLDLMDLITSLHDQGH 363
 Db 301 DQETEEVQOOLAPPFGHSAFAPEFOOTDSGKVLKQLARLDLMDLITSLHDQGH 360
 QY 364 GDP 366
 Db 361 GDP 363
 RESULT 10
 AAB90573
 ID AAB90573 standard; protein: 363 AA.
 AC AAB90573;
 XX 01-JUN-2001 (first entry)
 DT
 DE Human secreted protein, SEQ ID NO: 111.
 XX
 XX Human; secreted protein; immunomodulatory; antisclerotic;
 KM dermatological; anti-inflammatory; anti-HIV; cytostatic; cardiant;
 KM vascular; anti-angiogenic; ophthalmological; neuroprotectant;
 KM neurotropic; anticonvulsant; antialzheimers; antiparkinsonian;
 KM antimicrobial; vulnecary; vaccine; gene therapy; cancer;
 KM protein coordinate data; infection.
 XX
 OS Homo sapiens.
 PN WO200121658-A1.
 XX 29-MAR-2001.
 PD
 XX 22-SEP-2000; 2000WO-US26013.
 PF
 XX 24-SEP-1999; 99US-0155709.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI N.J. Baker KP, Birge CE, Ebner R, Fiscella M, Komatsoulis GA;
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
 PI Young PE, Wei P, Florence KA;
 XX WPI: 2001-235311/24.
 DR N-PSDB: AAF97913.
 XX
 PT Nucleic acids encoding 32 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's

PT disease and diabetic retinopathy -
 XX
 PS Claim 11; Page 794-795; 890pp; English.
 CC The present sequence is one of 32 novel human secreted polypeptides. The
 CC nucleic acid molecules and polypeptides may be used in the prevention,
 CC diagnosis and treatment of diseases such as immune disorders (e.g.
 CC multiple sclerosis, systemic lupus erythematosus and human
 CC immuno-deficiency virus (HIV) infections), hyperproliferative disorders
 CC (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. scleritis syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration
 CC and/or chemotaxis. The nucleic acid molecules may be used to produce the
 CC secreted polypeptides. They may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acid
 CC sequences in samples. The polypeptides may be used as antigens in the
 CC production of antibodies and in assays to identify modulators of
 CC their expression and activity.
 CC
 SQ Sequence 363 AA:

Query Match 99.0%; Score 1858; DB 22; Length 363;
 Best Local Similarity 99.7%; Pred. No. 1,8e-153;
 Matches 362; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MAAVLTWALALISAFSATOARKGFMDYFSQTSQDKGRVEQIHQOKMAREPATLKDSLEOD 63
 DB 1 MAAVLTWALALISAFSATOARKGFMDYFSQTSQDKGRVEQIHQOKMAREPATLKDSLEOD 60
 QY 64 LNNMNFLEKLRPI,SGSEAPRLPQDPVGMRRQLQFELFEVKARIQPYMAFAHELVGMNLE 123
 DB 61 LNNMNFLEKLRPI,SGSEAPRLPQDPVGMRRQLQFELFEVKARIQPYMAFAHELVGMNLE 120
 QY 124 GLRQOLKPYTMDLMEQVALRVQELQFQLRVSGEDTKAQLLGVDKEMALLQGLQSRVYHH 183
 DB 121 GLRQOLKPYTMDLMEQVALRVQELQFQLRVSGEDTKAQLLGVDKEMALLQGLQSRVYHH 180
 QY 184 TGRFKELFHPYAESLVSIGIRHVOELHRSVAPAPASPARLSKCYVLSRKLTLLKAKAH 243
 DB 181 TGRFKELFHPYAESLVSIGIRHVOELHRSVAPAPASPARLSKCYVLSRKLTLLKAKAH 240
 QY 244 ARIQOULDLREBLRAFAFGTGEBCAGDPQMLSEVQRQLQAFRODTYLIQIAAFTRAI 303
 DB 241 ARIQOULDLREBLRAFAFGTGEBCAGDPQMLSEVQRQLQAFRODTYLIQIAAFTRAI 300
 QY 304 DQETEEVOOQLAPPPGGHSAFAPEFOQTDSGVLSKQLQARLDLMDLWEDITHSLHDOGHSHL 363
 DB 301 DQETEEVOOQLAPPPGGHSAFAPEFOQTDSGVLSKQLQARLDLMDLWEDITHSLHDOGHSHL 360
 QY 364 GDP 366
 DB 361 GDP 363

RESULT 11
 ABC65475
 ID ABC65475 standard; Protein: 363 AA.
 XX
 AC ABG65475;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human albumin fusion protein #2150.
 XX
 KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cystostatic; antifertility; antiinflammatory; anticancer;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;

KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritis.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO20017137-A1.
 XX
 XX 18-OCT-2001.
 PD
 XX
 PD 12-APR-2001; 2001WO-US11988.
 PP
 PR 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Rosen CA, Haselaine WA;
 PI WPI; 2002-010886/01.
 DR
 XX
 XX
 PT New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein -
 PS Claim 1; Page 2047-2049; 2102pp; English.

The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or
 CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), haematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG6326-ABG6518 represent albumin
 CC fusion proteins of the invention.
 CC
 SQ Sequence 363 AA:

Query Match 99.0%; Score 1858; DB 23; Length 363;
 Best Local Similarity 99.7%; Pred. No. 1,8e-153;
 Matches 362; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MAAVLTWALALISAFSATOARKGFMDYFSQTSQDKGRVEQIHQOKMAREPATLKDSLEOD 63
 DB 1 MAAVLTWALALISAFSATOARKGFMDYFSQTSQDKGRVEQIHQOKMAREPATLKDSLEOD 60
 QY 64 LNNMNFLEKLRPI,SGSEAPRLPQDPVGMRRQLQFELFEVKARIQPYMAFAHELVGMNLE 123
 DB 61 LNNMNFLEKLRPI,SGSEAPRLPQDPVGMRRQLQFELFEVKARIQPYMAFAHELVGMNLE 120
 QY 124 GLRQOLKPYTMDLMEQVALRVQELQFQLRVSGEDTKAQLLGVDKEMALLQGLQSRVYHH 183
 DB 121 GLRQOLKPYTMDLMEQVALRVQELQFQLRVSGEDTKAQLLGVDKEMALLQGLQSRVYHH 180
 QY 184 TGRFKELFHPYAESLVSIGIRHVOELHRSVAPAPASPARLSKCYVLSRKLTLLKAKAH 243
 DB 181 TGRFKELFHPYAESLVSIGIRHVOELHRSVAPAPASPARLSKCYVLSRKLTLLKAKAH 240
 QY 244 ARIQOULDLREBLRAFAFGTGEBCAGDPQMLSEVQRQLQAFRODTYLIQIAAFTRAI 303
 DB 241 ARIQOULDLREBLRAFAFGTGEBCAGDPQMLSEVQRQLQAFRODTYLIQIAAFTRAI 300
 QY 304 DQETEEVOOQLAPPPGGHSAFAPEFOQTDSGVLSKQLQARLDLMDLWEDITHSLHDOGHSHL 363
 DB 301 DQETEEVOOQLAPPPGGHSAFAPEFOQTDSGVLSKQLQARLDLMDLWEDITHSLHDOGHSHL 360
 QY 364 GDP 366

DB 361 GDP 363

RESULT 12

AA086153

ID AA086153 standard; Protein; 274 AA.

XX AA086153:

15-JUL-2002 (first entry)

Human PRO773 polypeptide.

Human: PRO: benign tumour; malignant tumour; lymphoid malignancy;

leukemia; neuronal disorder; stromal disorder; blastocellular disorder;

inflammatory disorder; immune cytostatic; neuroprotective.

XX Homo sapiens.

PN WO200153486-A1.

26-JUL-2001.

11-FEB-2000; 2000WO-US03565.

08-MAR-1999; 99WO-US05028.

11-MAR-1999; 99US-123972P.

11-MAY-1999; 99US-133459P.

02-JUN-1999; 99WO-US12252.

22-JUN-1999; 99US-140650P.

20-JUL-1999; 99US-144758P.

26-JUL-1999; 99US-145698P.

28-JUL-1999; 99US-146222P.

17-AUG-1999; 99US-149395P.

31-AUG-1999; 99US-151689P.

01-SEP-1999; 99WO-US20111.

15-SEP-1999; 99WO-US21090.

30-NOV-1999; 99WO-US28313.

01-DEC-1999; 99WO-US28301.

01-DEC-1999; 99WO-US28634.

05-JAN-2000; 2000WO-US00219.

(GETH) GENENTECH INC.

XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;

PI Marsters SA, Pan J, Pilti KM, Roy MA, Smith V, Stone DM;

PI Watanabe CK, Wood WI;

DR WPI: 2002-205567/26.

DR N-PSDB; ABK40279.

XX Thirty five nucleic acids encoding PRO polypeptides, useful for

PT treating benign or malignant tumours, leukemias and lymphoid

PT malignancies, inflammatory, angiogenic and immunologic disorders -

PS Claim 61; Fig 52; 302pp; English.

Query Match 74.4%; Score 1396; DB 23; Length 274;
Best Local Similarity 100.0%; Pred. No. 2; 1e-113;
Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASHAAVLTWALALSAFSAATQARKGFMDYFSQTSQDKGRVEQIHQOKMAREPATLQSL 60

DB 1 MASHAAVLTWALALSAFSAATQARKGFMDYFSQTSQDKGRVEQIHQOKMAREPATLQSL 60

QY 61 EODLNNNNKFLKRLPLSGSEAPRLPDDPYGKRRQLODELEEVNARLQPTMAENHELVGM 120

DB 61 EODLNNNNKFLKRLPLSGSEAPRLPDDPYGKRRQLODELEEVNARLQPTMAENHELVGM 120

QY 121 NLEGLRQOLKPYTMDLMEQVALRVQLOEOLRVYGEDTKAQLGCVDEANALJQGLQSRV 180

DB 121 NLEGLRQOLKPYTMDLMEQVALRVQLOEOLRVYGEDTKAQLGCVDEANALJQGLQSRV 180

QY 181 VHHTRGKFLPHYAESLVSGIGHVQELHRSVAPHA PASPARLSKCVYLTKAK 240

DB 181 VHHTRGKFLPHYAESLVSGIGHVQELHRSVAPHA PASPARLSKCVYLTKAK 240

QY 241 ALHARIQNDQLREELSRFAAGTTEGAGPDP 274

DB 241 ALHARIQNDQLREELSRFAAGTTEGAGPDP 274

RESULT 13

ID ABG34037 standard; Protein; 274 AA.

ABG34037:

15-JUL-2002 (first entry)

Human Pro peptide #8.

Human; PHU, secreted protein; transmembrane protein;

genetic disorder; tumour; cancer.

XX Homo sapiens.

PN WO200224888-A2.

28-MAR-2002.

29-AUG-2001; 2001WO-US27099.

01-SEP-2000; 2000US-229896P.

05-SEP-2000; 2000US-230621P.

22-SEP-2000; 2000US-235147P.

10-NOV-2000; 2000WO-US30873.

12-JAN-2001; 2001US-261878P.

16-JAN-2001; 2001US-261910P.

16-JAN-2001; 2001US-261939P.

16-JAN-2001; 2001US-262150P.

25-JAN-2001; 2001US-264395P.

02-FEB-2001; 2001US-266421P.

09-FEB-2001; 2001US-267623P.

28-FEB-2001; 2001WO-US05520.

09-MAR-2001; 2001US-274399P.

03-APR-2001; 2001US-280982P.

04-APR-2001; 2001US-282129P.

09-MAY-2001; 2001US-290589P.

PR 25-MAY-2001; 2001WO-US17092.

PR 01-JUN-2001; 2001WO-US17600.

PR 20-JUN-2001; 2001WO-US19692.

PR 29-JUN-2001; 2001WO-US21066.

PR 09-JUL-2001; 2001WO-US21735.

(GETH) GENENTECH INC.

XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;

PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;

PI Fong S;
 XX
 DR WPI: 2002-362426/39.
 DR N-PSDB; ABR69968.
 XX
 PT New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or
 PT for genetic analysis of individuals with genetic disorders
 XX
 PS Claim 11; Figure 16; 218pp: English.

XX This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The
 CC invention also comprises a method for producing the proteins of the
 CC invention by recombinant means and antibodies specific for the protein
 CC of the invention. The antibody may be used for detecting the PRO
 CC proteins of the invention and may be used to modify their activity.
 CC polynucleotides may be used as hybridisation probes for a cDNA library
 CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
 CC construct hybridisation probes for mapping the gene which encodes that
 CC PRO and for genetic analysis of individuals with genetic disorders, in
 CC assays to identify other proteins or molecules involved in binding
 CC reaction, to generate transgenic animals or knock-out animals which in
 CC turn are useful in the development and screening of therapeutically
 CC useful reagents, for chromosome identification, and tissue typing. The
 CC PRO polypeptides are useful in gene therapy, and as molecular weight
 CC markers for protein electrophoresis purposes. The sequences may
 CC also be used to detect overexpression on PRO polypeptides in cancerous
 CC tumours and for screening for differentially expressed genes using
 CC microarray technology. The present sequence represents a human PRO
 CC protein of the invention.

XX Sequence 274 AA:

Query Match 73.6%; Score 1381; DB 23; Length 274;
 Best Local Similarity 99.3%; Pred. No. 4,4e-112;
 Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MASMAVILTMALALISAFSATOARKGFMDYFSOTSGDKGRVEQIHQOKMARREPATLKDSL 60
 DB 1 MASMAVILTMALALISAFSATOARKGFMDYFSOTSGDKGRVEQIHQOKMARREPATLKDSL 60
 QY 61 EDDLNNMFKLEKRLPSSSEAPRLPDYPVGMKRLQDELEEVKARLQPYMAEAHELVCVM 120
 DB 61 EDDLNNMFKLEKRLPSSSEAPRLPDYPVGMKRLQDELEEVKARLQPYMAEAHELVCVM 120
 QY 121 NIEGLRQOLKPYTMIDMEQVALRVQELQEQRLRVYGDHTKAOLICGVDFAMALLQGLQSRV 180
 DB 121 NIEGLRQOLKPYTMIDMEQVALRVQELQEQRLRVYGDHTKAOLICGVDFAMALLQGLQSRV 180
 QY 181 VHTGTFKELFHPYAKSLVSGICGRHVQELHRSVAHPAPASPARLSKCYQVLSNKLITLAKK 240
 DB 181 VHTGTFKELFHPYAKSLVSGICGRHVQELHRSVAHPAPASPARLSKCYQVLSNKLITLAKK 240
 QY 241 ALHARTIQNLQDLRELSRAFACTGTEGACGDP 274
 DB 241 ALHARTIQNLQDLRELSRAFACTGTEGACGDP 274

RESULT 14
 AAY54561
 ID AAY54561 standard; Protein: 367 AA.

XX AAY54561;
 AC
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Amino acid sequence of the rat RAP3 protein.
 XX
 KW RAP3; regeneration association protein 3; liver regeneration;
 XX liver proliferation.
 OS Rattus sp.

XX
 PN WO200003013-A2.
 XX
 PD 20-JAN-2000.
 XX
 PF 12-JUL-1999; 99MO-EP04938.
 XX
 PR 10-JUL-1998; 98EP-0202336.
 XX
 PA (AMST-) AMSTERDAM MOLECULAR THEKAPEUTICS NV.
 XX
 PI Chamulleau RAFM, Groenink M, Van Der Vliet HN, Leeuwater ACJ;
 XX
 DR WPI: 2000-147615/13.
 DR N-PSDB; AA425872, AA425873.
 XX
 PT Isolated RAP3 gene, protein and antibody useful for diagnosing liver
 PT regeneration and/or cell proliferation
 XX
 PS Claim 6; Fig 2; 42pp: English.

XX The present sequence represents a rat RAP3 (regeneration association
 CC protein 3). The RAP3 cDNA sequence was isolated from a rat liver cDNA
 CC library. The RAP3 gene is involved in regeneration processes of the
 CC liver. The RAP3 gene was found to be upregulated 6 hours after partial
 CC hepatectomy, after which it was downregulated. The RAP3 cDNA sequence
 CC is useful as a source for PCR primers and probes, which are useful
 CC for detecting nucleotide sequences in a source material. The RAP3
 CC sequence is useful as a marker of liver proliferation. The RAP3
 CC protein is useful for the diagnosis of liver regeneration and liver
 CC cell proliferation. RAP3 antibodies, PCR primers and probes are useful
 CC for detecting the occurrence of liver cell proliferation in a patient.
 CC The RAP3 protein is also useful for enhancing the growth of
 CC regeneration of liver tissue comprising treating the liver tissue
 CC such as extracorporeal or intracorporeal.

XX Sequence 367 AA:

Query Match 68.9%; Score 1293; DB 21; Length 367;
 Best Local Similarity 71.8%; Pred. No. 3e-104;
 Matches 260; Conservative 32; Mismatches 62; Indels 8; Gaps 4;

QY 4 MAVITMALALISAFSATOARKGFMDYFSOTSGDKGRVEQIHQOKMARREPATLKDSL 63
 DB 1 MAVITMALALISAFSATOARKGFMDYFSOTSGDKGRVEQIHQOKMARREPATLKDSL 63
 QY 64 LNNMFKLEKRLPSSSEAPRLPDYPVGMKRLQDELEEVKARLQPYMAEAHELVCVM 121
 DB 57 LNNMFKLEKRLPSSSEAPRLPDYPVGMKRLQDELEEVKARLQPYMAEAHELVCVM 116
 QY 122 LKELRQOLKPYTMIDMEQVALRVQELQEQRLRVYGDHTKAOLICGVDFAMALLQGLQSRV 181
 DB 117 LKELRQOLKPYTMIDMEQVALRVQELQEQRLRVYGDHTKAOLICGVDFAMALLQGLQSRV 176
 QY 182 HHTGTFKELFHPYAKSLVSGICGRHVQELHRSVAHPAPASPARLSKCYQVLSNKLITLAKK 241
 DB 177 HHTGTFKELFHPYAKSLVSGICGRHVQELHRSVAHPAPASPARLSKCYQVLSNKLITLAKK 236
 QY 242 LKELRQOLKPYTMIDMEQVALRVQELQEQRLRVYGDHTKAOLICGVDFAMALLQGLQSRV 299
 DB 237 LKELRQOLKPYTMIDMEQVALRVQELQEQRLRVYGDHTKAOLICGVDFAMALLQGLQSRV 296
 QY 300 TRALIDETEEVQOOLAPPFGSAFAPEFQGTGSCVLSKLQARLDDLMEDITHSLHDCG 359
 DB 297 TRALIDETEEVQOOLAPPFGSAFAPEFQGTGSCVLSKLQARLDDLMEDITHSLHDCG 356
 QY 360 HS 361
 DB 357 HS 358

RESULT 15
 AAY67399

ID AAV67399 standard; Protein: 367 AA.
 XX AAV67399;
 AC 25-APR-2000 (first entry)
 DT 25-APR-2000 (first entry)
 DE Rat rap3 protein sequence.
 XX Rat rap3 protein sequence.
 KM rap3; rat; liver regeneration; hepatic cell proliferation; liver biopsy;
 KW liver transplant; bioartificial liver.
 XX Rattus sp.
 OS Rattus sp.
 PN EP976824-A1.
 XX EP976824-A1.
 PD 02-FEB-2000.
 XX 02-FEB-2000.
 PF 10-JUL-1998; 98EP-02023336.
 XX 10-JUL-1998; 98EP-02023336.
 PR 10-JUL-1998; 98EP-02023336.
 XX 10-JUL-1998; 98EP-02023336.
 PA (AMST-) AMSTERDAM MOLECULAR THERAPEUTICS BV.
 XX (AMST-) AMSTERDAM MOLECULAR THERAPEUTICS BV.
 PI Chamuleau RAFM, Groenink M, Van der Vliet HN, Leegwater ACJ.
 XX Chamuleau RAFM, Groenink M, Van der Vliet HN, Leegwater ACJ.
 DR WPI: 2000-147615/13.
 DR N-PSDB; AAZ60187.
 XX N-PSDB; AAZ60187.
 PT Isolated RAP3 gene, protein and antibody useful for diagnosing liver
 PT regeneration and/or cell proliferation -
 PS Claim 6; Fig 2; 31pp; English.
 XX Claim 6; Fig 2; 31pp; English.
 CC This sequence represents the rat rap3 protein encoded by the RAP3 gene.
 CC The gene and protein are involved in the regeneration processes of the
 CC liver, and the RAP3 gene is expressed specifically in the liver. The RAP3
 CC gene is useful for designing PCR primers and probes for detecting
 CC nucleotide sequences in a source material and as a marker of liver
 CC proliferation. The rap3 protein is useful for the diagnosis of liver
 CC regeneration and/or liver cell proliferation. Anti-rap3 antibodies, PCR
 CC primers and nucleotide sequences which act as probes are useful for
 CC detecting the occurrence of liver cell proliferation in a patient. Single
 CC stranded oligonucleotides that are complementary to RAP3 can be used as
 CC probes to detect the amount of mRNA transcribed from RAP3 present in a
 CC sample such as a liver biopsy, plasma or serum of a tissue or body fluid
 CC in comparison to a reference sample. The probes can also be used for
 CC screening a liver cDNA or genomic library. The rap3 protein is useful for
 CC enhancing the growth or regeneration of liver tissue. The methods of the
 CC invention can be used to establish the efficacy of therapeutic agents
 CC stimulating liver regeneration and for patients who have undergone liver
 CC transplantation and for monitoring patients treated with a bioartificial
 CC liver.
 CC liver.
 XX liver.
 XX liver.
 SQ Sequence 367 AA;
 Query Match 68.9%; Score 1293; DB 21; Length 367;
 Best Local Similarity 71.8%; Pred. No. 3e-104;
 Matches 260; Conservative 32; Mismatches 62; Indels 8; Gaps 4;
 QY 4 NAAVTMALALISAFSAQARKGFWDFSGDKRVEQIHQOMAREPATLKDSEOD 63
 DB 1 NAAVTMALALISAFSAQARKGFWDFSGDKRVEQIHQOMAREPATLKDSEOD 56
 QY 64 INNNKKFLKRLPL--SGSEAPRLPDDPVGMRROLOEELEPYKARLOPYMAEHLVGMN 121
 DB 57 LYNNMNFLEKLGPLPEKPEPRRLADPEGIRKQLOQLELEPYSTRLEPYMAAKHQVGMN 116
 QY 122 LEGLRQOLKPYTMDLMEQVALRVEQLRVEGEDTKAQLLGVDAMALLGLQSRVY 181
 DB 117 LEGLRQOLKPYTMDLMEQVALRVEQLRVEGEDTKAQLLGVDAMALLGLQSRVY 176
 QY 182 HHTGFKELFHPYAEISVSGIRHVOELHRSVAPAPASPARLSRCVOYLSKRLTKAKA 241
 DB 177 HHTDVRKELFHPYAEISVSGIRHVOELHRSVAPAPASPARLSRCVOYLSKRLTKAKA 236
 QY 242 LHARTOONLQDLRELSR--AFACGTGEGAGPDPOMLSEVRORLOAFPRODTYLOIAAF 299
 DB 237 LHRTIQRLNLDQLRDELSTFIRVSTGADNRDSDLPQALSDVEVRORLOAFKHDTYLOIAAF 296
 QY 300 TRAIQDETEEVQOOLAPPPPGHSAFAPEFOQDSGKYLKLOARLDLMDITHSLHDG 359
 DB 297 TQALDQETEEIQHOLAPPSPHSAFAPELGHSDSNKALSRLOSLDLDLMDIAYGLHDG 356
 QY 360 HS 361
 DB 357 HS 358

Search completed: January 23, 2003, 17:43:48
 Job time : 63 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2003, 17:42:47 : Search time 22 Seconds
(without alignments)
489,490 Million cell updates/sec

Title: US-09-842-364-3

Perfect score: 1877
Sequence: 1 MASMAAVLTWALALALSAFSA.....LMEDITHSLHDGSHLGD 366

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCMTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1877	100.0	366	4	US-09-750-580-3 Sequence 3, Appli
2	200	10.7	267	1	US-07-959-946-3 Sequence 3, Appli
3	200	10.7	267	1	US-08-333-577-3 Sequence 3, Appli
4	200	10.7	267	4	US-08-952-796-2 Sequence 2, Appli
5	200	10.7	267	5	PCT-US92-08634-3 Sequence 3, Appli
6	183.5	9.8	264	1	US-08-448-606-6 Sequence 6, Appli
7	178	9.5	200	4	US-08-952-796-15 Sequence 15, Appli
8	162	8.6	317	4	US-08-949-155-6 Sequence 6, Appli
9	162	8.6	317	4	US-09-819-964-6 Sequence 6, Appli
10	149	7.9	317	1	US-07-709-949-2 Sequence 2, Appli
11	128.5	6.8	2101	1	US-08-456-380-4 Sequence 4, Appli
12	128.5	6.8	2101	1	US-08-470-950-4 Sequence 4, Appli
13	128.5	6.8	2101	1	US-08-467-781-4 Sequence 4, Appli
14	128.5	6.8	2101	1	US-08-195-487-4 Sequence 4, Appli
15	128.5	6.8	2101	2	US-08-483-924-4 Sequence 4, Appli
16	128.5	6.8	2101	2	US-09-452-294-1 Sequence 1, Appli
17	128.5	6.8	2101	5	PCT-US93-06160-4 Sequence 4, Appli
18	123	6.6	576	2	US-08-533-306A-2 Sequence 2, Appli
19	123	6.6	576	2	US-08-742-923A-2 Sequence 2, Appli
20	123	6.6	816	2	US-08-533-306A-6 Sequence 6, Appli
21	123	6.6	816	2	US-08-742-923A-6 Sequence 6, Appli
22	123	6.6	885	2	US-08-533-306A-4 Sequence 4, Appli
23	123	6.6	885	2	US-08-742-923A-4 Sequence 4, Appli
24	121.5	6.5	803	4	US-09-154-750A-85 Sequence 85, Appli
25	120	6.4	1618	1	US-07-853-913-4 Sequence 4, Appli
26	116	6.2	1388	4	US-09-572-191-2 Sequence 2, Appli
27	116	6.2	1388	4	US-09-723-262-2 Sequence 2, Appli

28	116	6.2	1388	4	US-09-723-219-2	Sequence 2, Appli
29	115	6.1	1090	4	US-09-085-1998-5	Sequence 5, Appli
30	115	6.1	1939	4	US-09-310-187A-1	Sequence 1, Appli
31	115	6.1	2293	4	US-09-368-590-2	Sequence 2, Appli
32	114.5	6.1	1886	4	US-08-938-105-3	Sequence 3, Appli
33	114	6.1	220	2	US-08-726-306A-29	Sequence 29, Appli
34	114	6.1	914	4	US-09-085-1998-4	Sequence 4, Appli
35	113	6.0	566	2	US-08-533-669A-8	Sequence 8, Appli
36	113	6.0	566	2	US-08-511-872-2	Sequence 2, Appli
37	113	6.0	566	4	US-09-183-861-8	Sequence 8, Appli
38	113	6.0	566	4	US-09-022-765-8	Sequence 8, Appli
39	113	6.0	2482	1	US-08-328-254-6	Sequence 6, Appli
40	112.5	6.0	1068	4	US-09-085-1998-11	Sequence 11, Appli
41	112	6.0	3248	5	US-08-353-700-1	Sequence 1, Appli
42	112	6.0	3248	5	PCT-US95-16216-1	Sequence 1, Appli
43	110	5.9	64	2	US-08-292-870-1	Sequence 1, Appli
44	107.5	5.7	331	4	US-09-695-458-2	Sequence 2, Appli
45	107.5	5.7	599	3	US-08-556-419-22	Sequence 22, Appli

ALIGNMENTS

```

RESULT 1
US-09-750-580-3
Sequence 3, Application US/09750580
Patent No. 6455280
GENERAL INFORMATION:
APPLICANT: Yen, Frances
APPLICANT: Denison, Blake
APPLICANT: Bour, Barbara
APPLICANT: Bihain, Bernard
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclert, Aymeric
APPLICANT: Bougueret, Lydie
APPLICANT: Ebbels-Reed, Dana
APPLICANT: Saller-Cid, Luisa
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
FILE REFERENCE: 89, US2, CIP
CURRENT APPLICATION NUMBER: US/09/750,580
CURRENT FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: US 09/599,362
PRIOR FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: PCT/IB00/0101
PRIOR FILING DATE: 2000-06-21
PRIOR APPLICATION NUMBER: PCT/IB99/02058
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: US 49/469/099
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: US 60/113,686
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: US 60/141,032
PRIOR FILING DATE: 1999-06-25
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent.pm
SEQ ID NO 3
LENGTH: 366
TYPE: PRT
ORGANISM: Homo sapiens
US-09-750-580-3
Query Match 100.0%; Score 1877; DB 4; Length 366;
Best Local Similarity 100.0%; Pred. No. 4.4e-161;
Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASMAAVLTWALALALSAFSAQARKGFWDYFSQSGDKGRVEQIHOQKMAREPTLKD 60
DB 1 MASMAAVLTWALALALSAFSAQARKGFWDYFSQSGDKGRVEQIHOQKMAREPTLKD 60
QY 61 EODLNNMKKFLKRLPSGSEAPRIPDPVGMRRQLOLELEEVYARQIOPYAEHLEL 120
DB 61 EODLNNMKKFLKRLPSGSEAPRIPDPVGMRRQLOLELEEVYARQIOPYAEHLEL 120

```

QY	121	NLEGI,ROOLKRYTMDLMEQVALRVQEOLEOOLRVVGEDTKKOLLGQVDEAMALLQGLDLSRV	180
Db	121	NLEGI,ROOLKRYTMDLMEQVALRVQEOLEOOLRVVGEDTKKOLLGQVDEAMALLQGLDLSRV	180
QY	181	VHHTGRFKELFHPVAESLVSGICGRHVOELHKSVAHPASPAPRLSRVCQVLSRKLTLKAK	240
Db	181	VHHTGRFKELFHPVAESLVSGICGRHVOELHKSVAHPASPAPRLSRVCQVLSRKLTLKAK	240
QY	241	ALHARIQONLDOLREELSRPAFGCTEGGAPDPDOMLSEFVROLOAFRODTYLOIAAFT	300
Db	241	ALHARIQONLDOLREELSRPAFGCTEGGAPDPDOMLSEFVROLOAFRODTYLOIAAFT	300
QY	301	RAIDQETEEVQOOLAPPBGHSAPAFEFQOTDSGKVLSQLQARLDLDMEDITITSLHDQGH	360
Db	301	RAIDQETEEVQOOLAPPBGHSAPAFEFQOTDSGKVLSQLQARLDLDMEDITITSLHDQGH	360
QY	361	SHLDGP	366
Db	361	SHLDGP	366

```

1      RESULT 2
2      US-07-959-946-3
3      Sequence 3, Application US/07959946
4      Patent No. 5408038
5      GENERAL INFORMATION:
6      APPLICANT: Smith, Richard K.
7      APPLICANT: Koduri, Raju
8      APPLICANT: Young, Stephen G.
9      APPLICANT: Wiltzong, Joseph L.
10     APPLICANT: Curtiss, Linda K.
11     TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
12     TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
13     NUMBER OF SEQUENCES: 20
14     CORRESPONDENCE ADDRESS:
15     ADDRESSEE: Dressler, Goldsmith, Shore, Sutter &
16     ADDRESSEE: Milamov, Ltd.
17     STREET: 180 No. 5408038th Sletson, Suite 4700
18     CITY: Chicago
19     STATE: Illinois
20     COUNTRY: USA
21     ZIP: 60601
22     COMPUTER READABLE FORM:
23     MEDIUM TYPE: Floppy disk
24     COMPUTER: IBM PC compatible
25     OPERATING SYSTEM: PC-DOS/MS-DOS
26     SOFTWARE: Patent In Release #1.0, Version #1.25
27     CURRENT APPLICATION DATA:
28     APPLICATION NUMBER: US/07/959,946
29     FILING DATE: 19921008
30     CLASSIFICATION: 435
31     PRIOR APPLICATION DATA:
32     APPLICATION NUMBER: US 07/901,706
33     FILING DATE: 18-JUN-1992
34     ATTORNEY/AGENT INFORMATION:
35     NAME: Gamson, Edward P.
36     REGISTRATION NUMBER: 29,381
37     REFERENCE/DOCKET NUMBER:
38     TELECOMMUNICATION INFORMATION:
39     TELEPHONE: (312)616-5400
40     TELEFAX: (312)616-5460
41     INFORMATION FOR SEQ ID NO: 3:
42     SEQUENCE CHARACTERISTICS:
43     LENGTH: 267 amino acids
44     TYPE: AMINO ACID
45     TOPOLOGY: linear
46     MOLECULE TYPE: protein
47     US-07-959-946-3

```

Query Match	10.7%;	Score 200;	DB 1;	Length 267;
Best Local Similarity	22.2%;	Pred. NO. 2.6e-10;		
Matches	74;	Conservative	61;	Mismatches 106;
			Indels	92;
			Gaps	10;

[illegible]

```

1      RESULT 3
2      US-08-333-577-3
3      : Sequence 3, Application US/08333577
4      : Patent No. 5786206
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Smith, Richard K.
8      : APPLICANT: Koduri, Raju
9      : APPLICANT: Young, Stephen G.
10     : APPLICANT: Witzium, Joseph L.
11     : APPLICANT: Curtiss, Linda K.
12     : TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
13     : TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
14     : NUMBER OF SEQUENCES: 20
15     :
16     : CORRESPONDENCE ADDRESSES:
17     : ADDRESSEE: Dressler, Goldsmith, Shore, Sultzer &
18     : ADDRESSEE: Milamov, Ltd.
19     : STREET: 180 No. 5786206th Stetson, Suite 4700
20     : CITY: Chicago
21     : STATE: Illinois
22     : COUNTRY: USA
23     :
24     : ZIP: 60601
25     :
26     : COMPUTER READABLE FORM:
27     : MEDIUM TYPE: Floppy disk
28     : COMPUTER: IBM PC compatible
29     : OPERATING SYSTEM: PC-DOS/MS-DOS
30     : SOFTWARE: Patent Release #1.0, Version #1.25
31     :
32     : CURRENT APPLICATION DATA:
33     : APPLICATION NUMBER: US/08/333,577
34     :
35     : FILING DATE:
36     : CLASSIFICATION: 530
37     :
38     : ATTORNEY/AGENT INFORMATION:
39     : NAME: Gansson, Edward P.
40     : REGISTRATION NUMBER: 29,381
41     : TELECOMMUNICATION INFORMATION: SCF 234.0
42     : TELEPHONE: (312)616-5400
43     : TELEFAX: (312)616-5460
44     : INFORMATION FOR SEQ ID NO: 3:
45     : SEQUENCE CHARACTERISTICS:
46     : LENGTH: 267 amino acids
47     : TYPE: amino acid
48     : TOPOLOGY: linear
49     :
50     : MOLECULE TYPE: protein
51     :
52     :
53     :
54     :
55     :
56     :
57     :
58     :
59     :
60     :
61     :
62     :
63     :
64     :
65     :
66     :
67     :
68     :
69     :
70     :
71     :
72     :
73     :
74     :
75     :
76     :
77     :
78     :
79     :
80     :
81     :
82     :
83     :
84     :
85     :
86     :
87     :
88     :
89     :
90     :
91     :
92     :
93     :
94     :
95     :
96     :
97     :
98     :
99     :
100    :
101    :
102    :
103    :
104    :
105    :
106    :
107    :
108    :
109    :
110    :
111    :
112    :
113    :
114    :
115    :
116    :
117    :
118    :
119    :
120    :
121    :
122    :
123    :
124    :
125    :
126    :
127    :
128    :
129    :
130    :
131    :
132    :
133    :
134    :
135    :
136    :
137    :
138    :
139    :
140    :
141    :
142    :
143    :
144    :
145    :
146    :
147    :
148    :
149    :
150    :
151    :
152    :
153    :
154    :
155    :
156    :
157    :
158    :
159    :
160    :
161    :
162    :
163    :
164    :
165    :
166    :
167    :
168    :
169    :
170    :
171    :
172    :
173    :
174    :
175    :
176    :
177    :
178    :
179    :
180    :
181    :
182    :
183    :
184    :
185    :
186    :
187    :
188    :
189    :
190    :
191    :
192    :
193    :
194    :
195    :
196    :
197    :
198    :
199    :
200    :
201    :
202    :
203    :
204    :
205    :
206    :
207    :
208    :
209    :
210    :
211    :
212    :
213    :
214    :
215    :
216    :
217    :
218    :
219    :
220    :
221    :
222    :
223    :
224    :
225    :
226    :
227    :
228    :
229    :
230    :
231    :
232    :
233    :
234    :
235    :
236    :
237    :
238    :
239    :
240    :
241    :
242    :
243    :
244    :
245    :
246    :
247    :
248    :
249    :
250    :
251    :
252    :
253    :
254    :
255    :
256    :
257    :
258    :
259    :
260    :
261    :
262    :
263    :
264    :
265    :
266    :
267    :
268    :
269    :
270    :
271    :
272    :
273    :
274    :
275    :
276    :
277    :
278    :
279    :
280    :
281    :
282    :
283    :
284    :
285    :
286    :
287    :
288    :
289    :
290    :
291    :
292    :
293    :
294    :
295    :
296    :
297    :
298    :
299    :
300    :
301    :
302    :
303    :
304    :
305    :
306    :
307    :
308    :
309    :
310    :
311    :
312    :
313    :
314    :
315    :
316    :
317    :
318    :
319    :
320    :
321    :
322    :
323    :
324    :
325    :
326    :
327    :
328    :
329    :
330    :
331    :
332    :
333    :
334    :
335    :
336    :
337    :
338    :
339    :
340    :
341    :
342    :
343    :
344    :
345    :
346    :
347    :
348    :
349    :
350    :
351    :
352    :
353    :
354    :
355    :
356    :
357    :
358    :
359    :
360    :
361    :
362    :
363    :
364    :
365    :
366    :
367    :
368    :
369    :
370    :
371    :
372    :
373    :
374    :
375    :
376    :
377    :
378    :
379    :
380    :
381    :
382    :
383    :
384    :
385    :
386    :
387    :
388    :
389    :
390    :
391    :
392    :
393    :
394    :
395    :
396    :
397    :
398    :
399    :
400    :
401    :
402    :
403    :
404    :
405    :
406    :
407    :
408    :
409    :
410    :
411    :
412    :
413    :
414    :
415    :
416    :
417    :
418    :
419    :
420    :
421    :
422    :
423    :
424    :
425    :
426    :
427    :
428    :
429    :
430    :
431    :
432    :
433    :
434    :
435    :
436    :
437    :
438    :
439    :
440    :
441    :
442    :
443    :
444    :
445    :
446    :
447    :
448    :
449    :
450    :
451    :
452    :
453    :
454    :
455    :
456    :
457    :
458    :
459    :
460    :
461    :
462    :
463    :
464    :
465    :
466    :
467    :
468    :
469    :
470    :
471    :
472    :
473    :
474    :
475    :
476    :
477    :
478    :
479    :
480    :
481    :
482    :
483    :
484    :
485    :
486    :
487    :
488    :
489    :
490    :
491    :
492    :
493    :
494    :
495    :
496    :
497    :
498    :
499    :
500    :
501    :
502    :
503    :
504    :
505    :
506    :
507    :
508    :
509    :
510    :
511    :
512    :
513    :
514    :
515    :
516    :
517    :
518    :
519    :
520    :
521    :
522    :
523    :
524    :
525    :
526    :
527    :
528    :
529    :
530    :
531    :
532    :
533    :
534    :
535    :
536    :
537    :
538    :
539    :
540    :
541    :
542    :
543    :
544    :
545    :
546    :
547    :
548    :
549    :
550    :
551    :
552    :
553    :
554    :
555    :
556    :
557    :
558    :
559    :
560    :
561    :
562    :
563    :
564    :
565    :
566    :
567    :
568    :
569    :
570    :
571    :
572    :
573    :
574    :
575    :
576    :
577    :
578    :
579    :
580    :
581    :
582    :
583    :
584    :
585    :
586    :
587    :
588    :
589    :
590    :
591    :
592    :
593    :
594    :
595    :
596    :
597    :
598    :
599    :
600    :
601    :
602    :
603    :
604    :
605    :
606    :
607    :
608    :
609    :
610    :
611    :
612    :
613    :
614    :
615    :
616    :
617    :
618    :
619    :
620    :
621    :
622    :
623    :
624    :
625    :
626    :
627    :
628    :
629    :
630    :
631    :
632    :
633    :
634    :
```

Query Match	10.7%;	Score 200;	DB 1;	Length 267;
Best Local Similarity	22.2%;	Pred. No. 2.6e-10;		
Matches 74;	Conservative 61;	Mismatches 106;	Indels 92;	Gaps 10;

```

OY 5 AAVLTWALALSAFSAATQARKGFWD-----YFSQTSQDGKR--VEQIH 45
    ||||| : : : ||| ||
Db 3 AAVLTWALVFL---TGSQARH-FWQODEPQSPMDRVKDLATVYVDVLDKSDRQVVSQFE 58
OY 46 OOKMARE-PATLKDSLLEODLNNMNFLEKLRPLSGSEAPRLPDQPVGMRRLOLELEEVK 104
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 GSALGKQNLNLKLDNWDVSTFSFKLRQDLGCVTQDFMDNLEKEETKEGLROEKSKDLEEVK 118
OY 105 ARLOPYMAVAHELVGMNLEGLRQDLKPYTMDLMEQVALRVQELQDLRVGSDTKAQLLG 164
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 AKVQPYLDDFOKKQWQEMELYRQKVEPLRAELQEGAROKLHLEQKLSPLGSEMR----- 173
OY 165 GVDEAMALLQGLQSRVHVHTGFKELFHPYAESLVSGIGRHVQELHRSVAPHPASPAPRL 224
    ||||| : : : ||| ||
Db 174 --DRARA-----HYDALRTHILAPYSDELRL 198
OY 225 SRCVOVLSRKLTKAKALHARIQNLIDLRELSAFAGTGEAGDPQMLSEVBQR 284
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 AARLEALKENGARLAEXHAKATEHLSTLSEKAKPAL-----EDLRQG 241
OY 285 ----LOAFKQDTYLOIAAFTRAIDQETEEVQOQ 313
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 LLPVLESEK-----VSFLSALLEYTKKLNTQ 267

```

RESULT 4

```

US-08-952-796-2
: Sequence 2, Application US/08952796
: Patent No. 6258596

```

GENERAL INFORMATION:

```

: APPLICANT: BENOTTI, Patrick
: APPLICANT: BRUCKERT, Eric
: APPLICANT: DENEFELE, Palrice
: APPLICANT: DUBERGER, Nicolas
: APPLICANT: FRUCHART, Jean-Charles
: APPLICANT: LUC, Gerald
: APPLICANT: TURPIN, Gerard
: APPLICANT: ASSMANN, Gerd
: APPLICANT: FUNKE, Harald
: TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-1
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Rhone-Poulenc Korer Inc.
: STREET: 500 Arcola Road, Mailstop 3C43
: CITY: Collegeville
: STATE: PA
: COUNTRY: USA

```

```

: ZIP: 19426
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/952.796
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: FR 95/06061
: FILING DATE: 22-MAY-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO FR96/00747
: FILING DATE: 20-MAY-1996

```

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Fehner Esq., Paul F.
: REGISTRATION NUMBER: 35,135
: REFERENCE/DOCKET NUMBER: ST95031-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (610) 454-3839
: TELEFAX: (610) 454-3808
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 267 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-952-796-2

```

```

Query Match 10.7%, Score 200, DB 4; length 267;
Best Local Similarity 22.2%, pred. No. 2,6c-10;
Matches 74; Conservative 61; Mismatches 106; Indels 92; Gaps 10;

```

```

OY 5 AAVLTWALALSAFSAATQARKGFWD-----YFSQTSQDGKR--VEQIH 45
    ||||| : : : ||| ||
Db 3 AAVLTWALVFL---TGSQARH-FWQODEPQSPMDRVKDLATVYVDVLDKSDRQVVSQFE 58
OY 46 OOKMARE-PATLKDSLLEODLNNMNFLEKLRPLSGSEAPRLPDQPVGMRRLOLELEEVK 104
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 GSALGKQNLNLKLDNWDVSTFSFKLRQDLGCVTQDFMDNLEKEETKEGLROEKSKDLEEVK 118
OY 105 ARLOPYMAVAHELVGMNLEGLRQDLKPYTMDLMEQVALRVQELQDLRVGSDTKAQLLG 164
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 AKVQPYLDDFOKKQWQEMELYRQKVEPLRAELQEGAROKLHLEQKLSPLGSEMR----- 173
OY 165 GVDEAMALLQGLQSRVHVHTGFKELFHPYAESLVSGIGRHVQELHRSVAPHPASPAPRL 224
    ||||| : : : ||| ||
Db 174 --DRARA-----HYDALRTHILAPYSDELRL 198
OY 225 SRCVOVLSRKLTKAKALHARIQNLIDLRELSAFAGTGEAGDPQMLSEVBQR 284
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 AARLEALKENGARLAEXHAKATEHLSTLSEKAKPAL-----EDLRQG 241
OY 285 ----LOAFKQDTYLOIAAFTRAIDQETEEVQOQ 313
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 LLPVLESEK-----VSFLSALLEYTKKLNTQ 267

```

RESULT 5

```

PCR-US92-08634-3

```

```

: Sequence 3, Application PC/TUS9208634
: GENERAL INFORMATION:

```

```

: APPLICANT: Smith, Richard K.
: APPLICANT: Koduri, Raju
: APPLICANT: Young, Stephen G.
: APPLICANT: Wiltz, Joseph L.
: APPLICANT: Curtis, Linda K.
: TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
: TITLE OF INVENTION: Pan Native Eptlope and Recombinant Antigens
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
: STREET: 180 North Sletson, Suite 4700
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA

```

```

: ZIP: 60601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/08634
: FILING DATE: 19921009
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/901,706
: FILING DATE: 18-JUN-1992

```

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Gamson, Edward P.
: REGISTRATION NUMBER: 29,381
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312)616-5400

```

TELEFAX: (312)616-5460
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 267 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US92-08634-3

Query Match 10.7%; Score 200; DB 5; length 267;
 Best local Similarity 22.2%; Pred. No. 2.6e-10;
 Matches 74; Conservative 61; Mismatches 106; Indels 92; Gaps 10;

QY 5 AAVLTWALLLAFSAFATQARKGFWD-----YFSQTSQDKGR--VEQI 45
 DB 3 AAVLTIAVLF--TGSQAKH-FWQDEPPQSPMDVKVCLATYVYVVKDSCGDVYSQF 58
 QY 46 QOKMARE-PATIKDSLEODLNMMKFEKLRPLSGSEARLPDPDPVGRROLOEELEEVK 104
 DB 59 GSAIGKQMLKLLDMDSVTSTFSKREQLGPVQEFWMNLKRETEGLRQEMSKDLEEVK 118
 QY 105 ARLQPMMAHFLVGMNLEGLRQOLKPYTMDLMEQVALVQELQRLVYVGEETKAQILG 164
 DB 119 AKVQPLDQFKRQWQEMELVYRQKVEPLRAELQEGARQKLIHELQKLSPLGEMR----- 173
 QY 165 GVDEAMALLQGLQSRVYHHTGFKELFHPYAESLYSGIRHVOELHRSVAPRHPAPAPARL 224
 DB 174 --DRARA-----HVALKTHLAPYSDELRLQRL 198
 QY 225 SRCVOVLSKRLKALKALHARIQONIDQRLBELSRFACTGTEEGAGPPQMLSEEVQR 284
 DB 199 ARLLEKLGKNGARLAEYHAKATEHLSTSEKAKPAL-----EDLRQ 241
 QY 285 ----LQAFRODTYLIQIAFPRAIDQETEVQOQ 313
 DB 242 LRPVLESFR-----VSFLSALTEETKRLNTO 267

RESULT 6
 US-08-448-606-6
 Sequence 6, Application US/08448606
 Patent No. 5721114

GENERAL INFORMATION:
 APPLICANT: Adrahams n, Lars
 APPLICANT: Holmgren, Erik
 APPLICANT: Kalden n, Christina
 APPLICANT: Lake, Mats
 APPLICANT: Mikaelsson, SA
 APPLICANT: Seijltz, Torsten
 TITLE OF INVENTION: Expression System For Producing
 TITLE OF SEQUENCES: Apolipoprotein AI-M
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pollock, Vande Sande & Priddy
 STREET: 1990 M Street, N.W., Suite 800
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/448,606
 FILING DATE: 25-AUG-1995
 CLASSIFICATION: 436
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/SE93/01061
 FILING DATE: 09-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9203753-0

FILING DATE: 11-DEC-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Amernick, Burton A.
 REGISTRATION NUMBER: 24,852
 REFERENCE/DOCKET NUMBER: 0151/00121
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)331-7111
 TELEFAX: (202) 293-6229
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 264 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-448-606-6

Query Match 9.8%; Score 183.5; DB 1; length 264;
 Best local Similarity 22.8%; Pred. No. 7.7e-09;
 Matches 62; Conservative 55; Mismatches 122; Indels 33; Gaps 6;

QY 1 MASMAVLTWALLA-LLSATSATQARKGFWD-----YFSQTSQDKGR--VEQI 50
 DB 1 MKRYTAIAIYVALAGFATVANADEPPQSPMDRVKIDATVYVDVVKDSQDVSQFEGSALG 60
 QY 51 RE-PATIKDSLEODLNMMKFEKLRPLSGSEARLPDPDPVGRROLOEELEEVKARLP 109
 DB 61 KQINIKLIDNDSDVTSTFSKREQLGPVQEFWMNLKRETEGLRQEMSKDLEEVKAVQ 120
 QY 110 YMAEAEHLVGMNLEGLRQOLKPYTMDLMEQVALVQELQRLVYVGEETKAQILG 169
 DB 121 YLDDQFKRQWQEMELVYRQKVEPLRAELQEGARQKLIHELQKLSPLGEMR-----DRA 173
 QY 170 WALQGLQSRVYHHTGFKELFHPYAESLYSGIRHVOELHRSVAPRHPAPAPARL 229
 DB 174 RAHVADLRTHLAPYSDELQCLAAKLEALKENGCARLAEYHAKATEHLSTSEKAKPAL 223
 QY 230 VLSRKLTKAKALHARIQONIDQRLBELSRFA 261
 DB 224 ----LSEKAKPALKDLRLQCLLPVLESFVSWF 250

RESULT 7
 US-08-952-796-15
 Sequence 15, Application US/08952796
 Patent No. 6258596

GENERAL INFORMATION:
 APPLICANT: BENNETT, Patrick
 APPLICANT: BRUCKERT, Eric
 APPLICANT: DENEFLE, Patrice
 APPLICANT: DUBERGER, Nicolas
 APPLICANT: FRUCHART, Jean-Charles
 APPLICANT: IUC, Gerard
 APPLICANT: TURPIN, Gerard
 APPLICANT: ASSMANN, Gerda
 APPLICANT: FUNKE, Harald
 TITLE OF INVENTION: NOVEL VARIANTS OF APOLIPOPROTEIN A-I
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, Mailstop 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/952,796
 FILING DATE:
 CLASSIFICATION: 424

```

1      APPLICATION NUMBER:  US 60/027,338
2      FILING DATE:  11-OCT-1996
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER:  US 60/046,094
5      FILING DATE:  09-MAY-1997
6      ATTORNEY/AGENT INFORMATION:
7      NAME:  Hibiier, David W.
8      REGISTRATION NUMBER:  41,071
9      REFERENCE/DOCKET NUMBER:  TAMK:177
10     TELECOMMUNICATION INFORMATION:
11     TELEPHONE:  (512) 418-3000
12     TELEFAX:  (713) 789-2679
13     INFORMATION FOR SEQ ID NO:  6:
14     SEQUENCE CHARACTERISTICS:
15     LENGTH:  317 amino acids
16     TYPE:  amino acid
17     TOPOLOGY:  linear
18     MOLECULE TYPE:  protein
19     US-08-949-155-6

```

Query Match	8.6%	Score	162	DB 4	Length	317			
Best Local Similarity	20.8%	Pred. No.	8.5e-07						
Matches	63	Conservative	66	Mismatches	112	Indels	62	Gaps	9

[illegible]

Db 292 QWA 294

RESULT 9
US-09-819-964-6
Sequence 6, Application US/09819964
Patent No. 6369294
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DUKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
A. APPLICATION NUMBER: US/09/819,964

FILING DATE: 28-Mar-2001
 CLASSIFICATION: <Unknown>
 APPLICATION NUMBER: 08/949,155
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 60/046,094
 FILING DATE: 09-MAY-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Hibler, David M.
 REGISTRATION NUMBER: 41,071
 REFERENCE/DOCKET NUMBER: TAMK:177
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (713) 789-2679
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 317 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 US-09-819-964-6

Query Match 8.6%; Score 162; DB 4; Length 317;
 Best Local Similarity 20.8%; Pred. No. 8.5e-07;
 Matches 63; Conservative 66; Mismatches 112; Indels 62; Gaps 9;

22 QARKGFMDYTSQSGDKGRV-EOIHQKMAREPATLKDSLEODLNKFKLEKLR---P 76
 45 QALGRFMDYLRWVQSLSDQVQELSTKYQE---LTELLESKEKKAIRRELEADLP 101
 77 LSSSEAPRLPDPPVGMKROLOELEEYKARLOPYMAEHLVGMNLEGLRQOQLPYTMDL 136
 102 VTETQARLSKELOAARAGADMEDVNRNLVYRSEVHNMLOQTEELSRSLASHIRKL 161
 137 MPQVALRVOELQOLRVYGEDTKAQLLQGVDEAMALIQOSRVVHHTGPKELFHPYAE 196
 162 RKKLLRTELQOKRLAVY---QAGLRKGRKRSVSLRERLGLPVEO-GRLR-----AA 210
 197 SLVSGIRHVOELHRSVAPAPASPARLSRCVOYLSRKLTLKAKALHARIQOINLQRLRE 256
 211 TLSTRAGQPLRER-----AAWGCOKLRGRLEEMGSRIRDRIDEKREQ 252
 257 LSAFAGTGEBCAGPDPQMLSEVQRLOAFQODPYLQIAATRAI---DOETEEVQO 312
 253 L-----BEVKYEEGQSQRLQAEFHALLKGFEPLEVEDIR 291
 313 QLA 315
 292 QWA 294

RESULT 10
 US-07-709-949-2
 Sequence 2, Application US/07709949
 Patent No. 5472858
 GENERAL INFORMATION:
 APPLICANT: Attie, Alan D.
 APPLICANT: Gretch, Daniel G.
 APPLICANT: Sturley, Stephen L.
 APPLICANT: Beckage, Nancy E.
 TITLE OF INVENTION: Production of Recombinant Proteins in
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Quares & Brady
 STREET: P.O. Box 2113
 CITY: Madison
 STATE: WI
 COUNTRY: USA
 ZIP: 53701-2113
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07709,949
 FILING DATE: 19910604
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Seay, Nicholas J.
 REGISTRATION NUMBER: 27,386
 REFERENCE/DOCKET NUMBER: 9629691801
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608) 251-5000
 TELEFAX: (608) 251-9166
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 317 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-709-949-2

Query Match 7.9%; Score 149; DB 1; Length 317;
 Best Local Similarity 22.0%; Pred. No. 1.3e-05;
 Matches 69; Conservative 59; Mismatches 117; Indels 68; Gaps 10;

10 MALALLSASFATQARKGFMDY--FSQTSQDKGRVEOIHQKMAREPATLKDSLEODLNK 67
 44 WELAL-----GRFMDYLRWVQSLSDQVQELSTKYQE---LTELLESKEKKAIRRELEADLP 92
 68 NKFL-EKLRPLSSSEAPRLPDPPVGMKROLOELEEYKARLOPYMAEHLVGMNLEGLR 126
 93 KSELEQTLTVAETRLARLSKELOTAQARLGADEMDVCGRLVYRGVQAMLGOSTEELR 152
 127 QQLKPYTMDLEQVALRVOELQOLRVYGEDTKAQLLQGVDEAMALIQOSRVVHHTG 186
 153 VLSASHRLKRLKRLRDPDLOKRLAVY---QAGLRKGRKRSVSLRERLGLPVEO-GRLR 201
 187 FKELFHPYAESLVSGIRHVOELHRSVAPAPASPARLSRCVOYLSRKLTLKAKALHARI 246
 202 LVEQGRVATVGSLSAQPLQER-----AAWGCOKLRGRLEEMGSRIRDRIDEKREQ 252
 247 QQNLIDQREELSRFACTGEBGAGPDPQMLSEVQRLOAFQODPYLQIAATRAI---DOETEEVQO 312
 244 RDRIDEVKEQVA-----EVRKLEEQOQIRLQAEAFQARLKS 282
 304 -DOETEEVQOQLA 315
 283 FHPYVHDMQKQWA 295

RESULT 11
 US-08-466-390-4
 Sequence 4, Application US/08466390
 Patent No. 5686562
 GENERAL INFORMATION:
 APPLICANT: TOURKATLY, GARY
 APPLICANT: LIDGARD, GRAHAM P.
 TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: TESTA, HURWITZ & THIBEAULT
 STREET: 125 HIGH STREET
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,390
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-390-4

```

```

Query Match 6.8%; Score 128.5; DB 1; Length 2101;
Best local Similarity 22.0%; Pred. No. 0.013;
Matches 81; Conservative 57; Mismatches 131; Indels 99; Gaps 15;

QY 37 DKGRVEQIHQOKMAREPATLKDSLEQDLNNKFLKRLPLSGSEAPRL----- 85
DB 453 ERGHFEERKQQL-----SLITDLOSSISNLSQAKKELEQASQAHGARTLNOVASLTSEL 507
QY 86 -----PODPVGMROLOEELEEVKARLOPYMAEAHVLGYNLJGRLROOLKPYTMD 135
DB 508 TTLNATIQOQDDELGLKQAKKEKQOLAQTLOQOEOAS-----QGLRHVQELSSS 559
QY 136 L--MEQVALRYVOELQ-----QLRVGEDTAKOGLGVDEAMALIGIOS-----R 179
DB 560 LKQKQOOLKEVAEKOQEAOTRODHAQOLATAEERESLSRER--DAALKQLEALEKEKAALKE 618
QY 180 VVHTGRKELFHPYAESLVSGIGHVOELHRSVAAPHASPAPSLRSCV----- 228
DB 619 ILQOQLOVANBARSDAOTSVTQAKREKALSRY-----EELQACVETAROBQHEAQ 670
QY 229 -QVLSRLTLTKAKALHA-----RIOONI.DQLELSRAFACTGTEGACPDPMI,SEEVRO 283
DB 671 AQVALELQQLSEQOKATEKERVAKQEKDQLOALKESELKVTGKS-----LEEKR 724
QY 284 RLQAF--RODTYLOIAAFTRAIQDETEVQOQLAPPPGSAFAPEFOOTDSGVLSKIQ 341
DB 725 AADALEEQORCISELKAETRSL-----VEQ-----HKRERKELEERAGR--KGLE 768
QY 342 ARLDLME 349
DB 769 ARLLQGE 776

```

```

RESULT 12
US-08-470-950-4
Sequence 4, Application US/08470950
Patent No. 5698439
GENERAL INFORMATION:
APPLICANT: TONKATLY, GARY
APPLICANT: LIDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
NUMBER OF SEQUENCES: 6
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,950
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-470-950-4

```

```

Query Match 6.8%; Score 128.5; DB 1; Length 2101;
Best local Similarity 22.0%; Pred. No. 0.013;
Matches 81; Conservative 57; Mismatches 131; Indels 99; Gaps 15;

QY 37 DKGRVEQIHQOKMAREPATLKDSLEQDLNNKFLKRLPLSGSEAPRL----- 85
DB 453 ERGHFEERKQQL-----SLITDLOSSISNLSQAKKELEQASQAHGARTLNOVASLTSEL 507
QY 86 -----PODPVGMROLOEELEEVKARLOPYMAEAHVLGYNLJGRLROOLKPYTMD 135
DB 508 TTLNATIQOQDDELGLKQAKKEKQOLAQTLOQOEOAS-----QGLRHVQELSSS 559
QY 136 L--MEQVALRYVOELQ-----QLRVGEDTAKOGLGVDEAMALIGIOS-----R 179
DB 560 LKQKQOOLKEVAEKOQEAOTRODHAQOLATAEERESLSRER--DAALKQLEALEKEKAALKE 618
QY 180 VVHTGRKELFHPYAESLVSGIGHVOELHRSVAAPHASPAPSLRSCV----- 228
DB 619 ILQOQLOVANBARSDAOTSVTQAKREKALSRY-----EELQACVETAROBQHEAQ 670
QY 229 -QVLSRLTLTKAKALHA-----RIOONI.DQLELSRAFACTGTEGACPDPMI,SEEVRO 283
DB 671 AQVALELQQLSEQOKATEKERVAKQEKDQLOALKESELKVTGKS-----LEEKR 724
QY 284 RLQAF--RODTYLOIAAFTRAIQDETEVQOQLAPPPGSAFAPEFOOTDSGVLSKIQ 341
DB 725 AADALEEQORCISELKAETRSL-----VEQ-----HKRERKELEERAGR--KGLE 768
QY 342 ARLDLME 349
DB 769 ARLLQGE 776

```

```

RESULT 13
US-08-467-781-4
Sequence 4, Application US/08467781
Patent No. 5780596
GENERAL INFORMATION:
APPLICANT: TONKATLY, GARY
APPLICANT: LIDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
NUMBER OF SEQUENCES: 6
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/467,781
? FILING DATE: 06-JUN-1995
? CLASSIFICATION: 424
? ATTORNEY/AGENT INFORMATION:
? NAME: PITCHER ESO, EDMUND R
? REGISTRATION NUMBER: 27,829
? REFERENCE/DOCKET NUMBER: MTP-013
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 248-7000
? TELEFAX: (617) 248-7100
? INFORMATION FOR SEQ. ID NO.: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2101 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? OS-08-467-781-4

```

```

COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,924
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
FAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2101 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-924-4

```

```

Query Match          6.8%; Score 128.5; DB 2; Length 2101;
Best Local Similarity 22.0%; Pred. No. 0.013;
Matches 81; Conservative 57; Mismatches 131; Indels 99; Gaps 15;

```

```

QY 37 DKGRVQIHQOMAREPATLKDSEODJNNMKNFKLEKPLUSGSPAPRI----- 85
Db 453 ERGHFEERKQDL-----SLITDLOSISINLSQAKEELQASQAGARLTAAVASLTSEL 507
QY 86 -----PODPVGMFRUQOELEEVKARLQPYMAEAHVELVGNLGLRQOLKPYTMD 135
Db 508 TTLNATIQOQDDELAGLKQAKKQAOQLAQTIQOQEOAS-----QGLRHQVEQI,SSS 559
QY 136 L--MEQVALRYOELQ-----QLRVGEDTKAQLLGVDAMALLQGLQS-----R 179
Db 560 LKQEQQLKEVAEKEATRODIAQOLATAEEREASLRER-DALQLQLEALEKEKAKLE 618
QY 180 VVHHTGRKRELPHRYAESIVSGICRHVOELHKSVAHPASPAPRLSRCV----- 228
Db 619 ILQOQLOVANEARDSAQSVTQAOEREKAELSRKV-----EELQACVETARQDQHEAQ 670
QY 229 -QVLSRKLTLKAKALIA---RIQONLDOLREELSRFAAGTGEAGFDPOMLSEEVRO 283
Db 671 AQVAELFIQLKSEQOKATEKKHVAQEKQDQLQKQALKSLKVTGKS-----LFEKKR 724
QY 284 RLQAF--KQDTYLOIAFTRAIDQETEEVQOQLAPPFGHSFAPEFQUTDSGKVLKQL 341
Db 725 AADALEEQRCISELKAEIRSL-----VEQ-----HKRERKLEEBRAGR--KGLE 768
QY 342 ARLDLME 349
Db 769 ARLDLGE 776

```

```

Search completed: January 23, 2003, 17:47:23
Job time : 25 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2003, 17:43:52 : Search time 40 seconds

(without alignments)
184.634 Million cell updates/sec

Title: US-09-842-364-3

Perfect score: 1877

Sequence: 1 MASMAVLTALALLSAFSA.....LMEDITHSLDGHSLGDP 366

Scoring table: HSIOSUM62

GAPOD 10.0 , GAPOD 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1877	100.0	366	10	US-09-751-877-3
2	1877	100.0	366	10	US-09-835-996A-2
3	1877	100.0	400	10	US-09-835-996A-41
4	1871	99.7	366	10	US-09-800-729-212
5	1858	99.0	363	10	US-09-800-729-111
6	350	18.6	382	9	US-09-987-107-37
7	350	18.6	382	10	US-09-800-729-206
8	333	17.7	429	9	US-09-987-107-34
9	328	17.5	396	10	US-09-800-729-207
10	327	17.4	396	9	US-09-987-107-33
11	315	16.8	401	9	US-09-987-107-36
12	311.5	16.6	391	9	US-09-987-107-36
13	298.5	15.9	391	10	US-09-800-729-208
14	298.5	15.9	395	9	US-09-987-107-35
15	200	10.7	267	9	US-09-803-918A-2
16	200	10.7	267	10	US-09-803-918A-2
17	193	10.3	267	9	US-09-987-107-16
18	189	10.1	258	9	US-09-987-107-4
19	189	10.1	273	9	US-09-987-107-50

20	188	10.0	301	9	US-09-987-107-3	Sequence 3, Appl1
21	188	10.0	301	9	US-09-987-107-5	Sequence 5, Appl1
22	188	10.0	316	9	US-09-987-107-48	Sequence 48, Appl1
23	188	10.0	316	9	US-09-987-107-54	Sequence 54, Appl1
24	186	9.9	304	9	US-09-987-107-6	Sequence 6, Appl1
25	186	9.9	304	9	US-09-987-107-7	Sequence 7, Appl1
26	186	9.9	304	9	US-09-987-107-8	Sequence 8, Appl1
27	186	9.9	323	9	US-09-987-107-56	Sequence 56, Appl1
28	186	9.9	323	9	US-09-987-107-58	Sequence 58, Appl1
29	186	9.9	323	9	US-09-987-107-60	Sequence 60, Appl1
30	185.5	9.9	243	9	US-09-987-107-1	Sequence 1, Appl1
31	185.5	9.9	243	9	US-09-987-107-2	Sequence 2, Appl1
32	185.5	9.9	261	9	US-09-987-107-52	Sequence 52, Appl1
33	185.5	9.9	306	9	US-09-987-107-9	Sequence 9, Appl1
34	185.5	9.9	306	9	US-09-987-107-10	Sequence 10, Appl1
35	185.5	9.9	306	9	US-09-987-107-11	Sequence 11, Appl1
36	185.5	9.9	324	9	US-09-987-107-62	Sequence 62, Appl1
37	185.5	9.9	324	9	US-09-987-107-64	Sequence 64, Appl1
38	185.5	9.9	324	9	US-09-987-107-66	Sequence 66, Appl1
39	185.5	9.9	329	9	US-09-987-107-14	Sequence 14, Appl1
40	185.5	9.9	336	9	US-09-987-107-44	Sequence 44, Appl1
41	185.5	9.9	337	9	US-09-987-107-46	Sequence 46, Appl1
42	185.5	9.9	344	9	US-09-987-107-68	Sequence 68, Appl1
43	185	9.9	266	9	US-09-987-107-20	Sequence 20, Appl1
44	183	9.7	264	9	US-09-987-107-25	Sequence 25, Appl1
45	176	9.4	265	9	US-09-987-107-17	Sequence 17, Appl1

ALIGNMENTS

RESULT 1	US-09-751-877-3	Sequence 3, Appl1	US/09751877
Patent No.	US20020142949A1	GENERAL INFORMATION:	
APPLICANT:	Denison, Blake	APPLICANT:	Yen, Frances
APPLICANT:	Bout, Barbara	APPLICANT:	Bihain, Bernard
APPLICANT:	Dumas Mline Edwards, Jean-Baptiste	APPLICANT:	Duclet, Aymeric
APPLICANT:	Bougueret, Lydie	APPLICANT:	Ebbels-Red, Dana
APPLICANT:	Salter-Cid, Luisa	TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH	
FILE REFERENCE:	89, US3, RRG	CURRENT APPLICATION NUMBER:	US/09751, 877
CURRENT FILING DATE:	2000-12-28	NUMBER OF SEQ ID NOS:	6
SOFTWARE:	Patent, pm	SEQ ID NO 3	
LENGTH:	366	TYPE:	PR
ORGANISM:	Homo sapiens	Query Match	100.0%, Score 1877; DB 10; Length 366;
Best Local Similarity	100.0%, Pred. No. 1,1e-124;	Matches 366; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	MASMAVLTALALLSAFSA	QY	1 MASMAVLTALALLSAFSA
1	EDLNMMNFLEKRLPSSSEAPRLPQDPVGNRQLOEELVYKARLPYMAEHLVGM	QY	1 EDLNMMNFLEKRLPSSSEAPRLPQDPVGNRQLOEELVYKARLPYMAEHLVGM
61	EDLNMMNFLEKRLPSSSEAPRLPQDPVGNRQLOEELVYKARLPYMAEHLVGM	QY	61 EDLNMMNFLEKRLPSSSEAPRLPQDPVGNRQLOEELVYKARLPYMAEHLVGM
121	NEGLRQOLKPYTMDLMEQVALRVDELQRLRVGEGDTKAQLLGVDAMALLGIGSRV	QY	121 NEGLRQOLKPYTMDLMEQVALRVDELQRLRVGEGDTKAQLLGVDAMALLGIGSRV
121	NEGLRQOLKPYTMDLMEQVALRVDELQRLRVGEGDTKAQLLGVDAMALLGIGSRV	QY	121 NEGLRQOLKPYTMDLMEQVALRVDELQRLRVGEGDTKAQLLGVDAMALLGIGSRV

```

OY 181 VHHTRKRELEFHPYAEELVSGIGRHVOELHRSVAPHAPASPARLSRCVYLSRKLTLKAK 240
    |||||||
DB 181 VHHTRKRELEFHPYAEELVSGIGRHVOELHRSVAPHAPASPARLSRCVYLSRKLTLKAK 240
OY 241 ALHARIQONLDQRELSRAFACTGTEEGAGPDPOMLSEEVROLOAFRODTYLOIAAFT 300
    |||||||
DB 241 ALHARIQONLDQRELSRAFACTGTEEGAGPDPOMLSEEVROLOAFRODTYLOIAAFT 300
OY 301 RAIDQETEEVQOOLAPPHPHSAPAFPEFOOTDSCKVLSKLOARLDLMDITSLHDOGH 360
    |||||||
DB 301 RAIDQETEEVQOOLAPPHPHSAPAFPEFOOTDSCKVLSKLOARLDLMDITSLHDOGH 360
OY 361 SHLGDP 366
    |||||
DB 361 SHLGDP 366

```

```

RESULT 2
US-09-835-996A-2
: Sequence 2, Application US/09835996A
: Patent No. US20020142953A1
: GENERAL INFORMATION:
: APPLICANT: Ballinger, Dennis
: APPLICANT: Loeb, Debra
: APPLICANT: Montgomery, Julie
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhao, Qiong
: APPLICANT: Weinman, Tom
: APPLICANT: Dirmann, Radoje
: APPLICANT: Ren, Feiyen
: APPLICANT: Qian, Xiahong
: APPLICANT: Wang, Duanli
: TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
: FILE REFERENCE: 28110/35915A
: CURRENT APPLICATION NUMBER: US/09/835, 996A
: CURRENT FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: US 60/197,137
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: US 09/714,936
: PRIOR FILING DATE: 2000-11-17
: PRIOR APPLICATION NUMBER: US 09/667,298
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: US 09/631,451
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 09/598,042
: PRIOR FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 366
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-835-996A-2

Query Match      100.0%; Score 1877; DB 10; Length 366;
Best Local Similarity 100.0%; Pred. No. 1,1e-124;
Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 181 VHHTRKRELEFHPYAEELVSGIGRHVOELHRSVAPHAPASPARLSRCVYLSRKLTLKAK 240
    |||||||
DB 181 VHHTRKRELEFHPYAEELVSGIGRHVOELHRSVAPHAPASPARLSRCVYLSRKLTLKAK 240
OY 241 ALHARIQONLDQRELSRAFACTGTEEGAGPDPOMLSEEVROLOAFRODTYLOIAAFT 300
    |||||||
DB 241 ALHARIQONLDQRELSRAFACTGTEEGAGPDPOMLSEEVROLOAFRODTYLOIAAFT 300
OY 301 RAIDQETEEVQOOLAPPHPHSAPAFPEFOOTDSCKVLSKLOARLDLMDITSLHDOGH 360
    |||||||
DB 301 RAIDQETEEVQOOLAPPHPHSAPAFPEFOOTDSCKVLSKLOARLDLMDITSLHDOGH 360
OY 361 SHLGDP 366
    |||||
DB 361 SHLGDP 366

```

```

RESULT 3
US-09-835-996A-41
: Sequence 41, Application US/09835996A
: Patent No. US20020142953A1
: GENERAL INFORMATION:
: APPLICANT: Ballinger, Dennis
: APPLICANT: Loeb, Debra
: APPLICANT: Montgomery, Julie
: APPLICANT: Tang, Y. Tom
: APPLICANT: Zhou, Ping
: APPLICANT: Goodrich, Ryle
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhao, Qiong
: APPLICANT: Weinman, Tom
: APPLICANT: Dirmann, Radoje
: APPLICANT: Ren, Feiyen
: APPLICANT: Qian, Xiahong
: APPLICANT: Wang, Duanli
: TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
: FILE REFERENCE: 28110/35915A
: CURRENT APPLICATION NUMBER: US/09/835, 996A
: CURRENT FILING DATE: 2001-04-16
: PRIOR APPLICATION NUMBER: US 60/197,137
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: US 09/714,936
: PRIOR FILING DATE: 2000-11-17
: PRIOR APPLICATION NUMBER: US 09/667,298
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: US 09/631,451
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 09/598,042
: PRIOR FILING DATE: 2000-06-20
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 41
: LENGTH: 400
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-835-996A-41

Query Match      100.0%; Score 1877; DB 10; Length 400;
Best Local Similarity 100.0%; Pred. No. 1,2e-124;
Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 155 NLEGLRQOLKPYTMDLMEQVALRVOELQBOILRVGDEDTKAQLGCVDEAMALLQGLQSRV 214

QY 181 VHTTGRKELFHPYAESLVSGIGHVOELHRSVAPHA PASPARLSRCVQVLSRRLTKAK 240

Db 215 VHTTGRKELFHPYAESLVSGIGHVOELHRSVAPHA PASPARLSRCVQVLSRRLTKAK 274

QY 241 ALHARLQONLDOLREELSRFAGTCEGAGPPDQMLSEEVROQLQAFRODTYLOIAAFT 300

Db 275 ALHARLQONLDOLREELSRFAGTCEGAGPPDQMLSEEVROQLQAFRODTYLOIAAFT 334

QY 301 RAIDQETEEVQOOLAPPFGHSAFAPEFQOTDSGKVL SKLQARLDLMDITTHSLHDQGH 360

Db 335 RAIDQETEEVQOOLAPPFGHSAFAPEFQOTDSGKVL SKLQARLDLMDITTHSLHDQGH 394

QY 361 SHLQDP 366

Db 395 SHLQDP 400

RESULT 4

US-09-800-729-212

Sequence 212, Application US/09800729

Patent No. US20020068319A1

GENERAL INFORMATION:

APPLICANT: Ni et al.

TITLE OR INVENTION: 32 Human secreted proteins

FILE REFERENCE: P2044P1

CURRENT APPLICATION NUMBER: US/09/800,729

CURRENT FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: PCT/US00/26013

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: 60/155,709

PRIOR FILING DATE: 1999-09-24

NUMBER OF SEQ ID NOS: 217

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 212

LENGTH: 366

TYPE: PRT

ORGANISM: Homo sapiens

US-09-800-729-212

Query Match

Best Local Similarity 99.7%; Score 1871; DB 10; Length 366;

Matches 365; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MASMAAVLTWALALLSFSATQARKGFWDYFSQTSQDKGRVEQIHQOKMAREPATLKDLSL 60

Db 1 MASMAAVLTWALALLSFSATQARKGFWDYFSQTSQDKGRVEQIHQOKMAREPATLKDLSL 60

QY 61 EODLNNKFLFKLRPLSGSPAPRLPQDPVGMROLOPELLEVKARIQPYWAEHVLVGM 120

Db 61 EODLNNKFLFKLRPLSGSPAPRLPQDPVGMROLOPELLEVKARIQPYWAEHVLVGM 120

QY 121 NLEGLRQOLKPYTMDLMEQVALRVOELQBOILRVGDEDTKAQLGCVDEAMALLQGLQSRV 180

Db 121 NLEGLRQOLKPYTMDLMEQVALRVOELQBOILRVGDEDTKAQLGCVDEAMALLQGLQSRV 180

QY 181 VHTTGRKELFHPYAESLVSGIGHVOELHRSVAPHA PASPARLSRCVQVLSRRLTKAK 240

Db 181 VHTTGRKELFHPYAESLVSGIGHVOELHRSVAPHA PASPARLSRCVQVLSRRLTKAK 240

QY 241 ALHARLQONLDOLREELSRFAGTCEGAGPPDQMLSEEVROQLQAFRODTYLOIAAFT 300

Db 241 ALHARLQONLDOLREELSRFAGTCEGAGPPDQMLSEEVROQLQAFRODTYLOIAAFT 300

QY 301 RAIDQETEEVQOOLAPPFGHSAFAPEFQOTDSGKVL SKLQARLDLMDITTHSLHDQGH 360

Db 301 RAIDQETEEVQOOLAPPFGHSAFAPEFQOTDSGKVL SKLQARLDLMDITTHSLHDQGH 360

QY 361 SHLQDP 366

Db 361 SHLQDP 366

RESULT 5

US-09-800-729-111

Sequence 111, Application US/09800729

Patent No. US20020068319A1

GENERAL INFORMATION:

APPLICANT: Ni et al.

TITLE OR INVENTION: 32 Human secreted proteins

FILE REFERENCE: P2044P1

CURRENT APPLICATION NUMBER: US/09/800,729

CURRENT FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: PCT/US00/26013

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: 60/155,709

PRIOR FILING DATE: 1999-09-24

NUMBER OF SEQ ID NOS: 217

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 111

LENGTH: 363

TYPE: PRT

ORGANISM: Homo sapiens

US-09-800-729-111

Query Match

Best Local Similarity 99.0%; Score 1858; DB 10; Length 363;

Matches 362; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MAAVLTWALALLSFSATQARKGFWDYFSQTSQDKGRVEQIHQOKMAREPATLKDLSLEOD 63

Db 1 MAAVLTWALALLSFSATQARKGFWDYFSQTSQDKGRVEQIHQOKMAREPATLKDLSLEOD 60

QY 64 LNNNKKFLFKLRPLSGSPAPRLPQDPVGMROLOPELLEVKARIQPYWAEHVLVGM 123

Db 61 LNNNKKFLFKLRPLSGSPAPRLPQDPVGMROLOPELLEVKARIQPYWAEHVLVGM 120

QY 124 GLRQOLKPYTMDLMEQVALRVOELQBOILRVGDEDTKAQLGCVDEAMALLQGLQSRV 183

Db 121 GLRQOLKPYTMDLMEQVALRVOELQBOILRVGDEDTKAQLGCVDEAMALLQGLQSRV 180

QY 184 TGRKELFHPYAESLVSGIGHVOELHRSVAPHA PASPARLSRCVQVLSRRLTKAKALH 243

Db 181 TGRKELFHPYAESLVSGIGHVOELHRSVAPHA PASPARLSRCVQVLSRRLTKAKALH 240

QY 244 ARIQONLDOLREELSRFAGTCEGAGPPDQMLSEEVROQLQAFRODTYLOIAAFTRAI 303

Db 241 ARIQONLDOLREELSRFAGTCEGAGPPDQMLSEEVROQLQAFRODTYLOIAAFTRAI 300

QY 304 DOETEEVQOOLAPPFGHSAFAPEFQOTDSGKVL SKLQARLDLMDITTHSLHDQGHSHL 363

Db 301 DOETEEVQOOLAPPFGHSAFAPEFQOTDSGKVL SKLQARLDLMDITTHSLHDQGHSHL 360

QY 364 GDP 366

Db 361 GDP 363

RESULT 6

US-09-987-107-37

Sequence 37, Application US/0987107

Patent No. US20020156007A1

GENERAL INFORMATION:

APPLICANT: GRAVERSEN, Jonas

TITLE OR INVENTION: APOLOPROTEINS ANALOGUES

FILE REFERENCE: GRAVERSENIA

CURRENT APPLICATION NUMBER: US/09/987,107

CURRENT FILING DATE: 2001-11-13

PRIOR APPLICATION NUMBER: US 60/264,022

PRIOR FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: DK PA2001 00057

PRIOR FILING DATE: 2001-01-15

PRIOR APPLICATION NUMBER: DK PA2000 01682

PRIOR FILING DATE: 2000-11-10

RESULT 9
US-09-800-729-207
; Sequence 207, Application US/09800729
; Patent No. US20020068319A1


```

: GENERAL INFORMATION:
: APPLICANT: Ni et al.
: TITLE OF INVENTION: 32 Human secreted proteins
: FILE REFERENCE: P2044P1
: CURRENT APPLICATION NUMBER: US/09/800,729
: CURRENT FILING DATE: 2001-03-08
: PRIOR APPLICATION NUMBER: PCT/US00/26013
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: 60/155,709
: PRIOR FILING DATE: 1999-09-24
: NUMBER OF SEQ ID NOS: 217
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 207
: LENGTH: 396
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-800-729-207

```

```

Query Match      17.5%; Score 328; DB 10; Length 396;
Best Local Similarity 28.4%; Pred. No. 4.5e-16;
Matches 96; Conservative 72; Mismatches 152; Indels 18; Gaps 6;

```

```

QY 5 AAVLTALALLSA-----FSATQARKGFMDYFSQTSQD-KGRVEDIHQOKMARE-PATLKD 58
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 5 AAVLTALVAAGARAEVSADQATVVMYDFSOI.SNNKAEVHLQKSEI.TQOLNALPHOD 64
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 59 SLEQDNLNMNKKFLEKLRPLSGSEAPRLPDPVGMRRQLOEELFEVKARLQPYMAEAHEL 118
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 65 KLGSEVNTYAGDLOKLVFPFTELHERLAKNSEKLEIGKELEELRRLPLPHANEVSQKI 124
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 119 GWNLEGIROOLKPYTMDLMEQVALRVQEIQYQOLRVGCEDTKAQILGCVDAAMALI.Q---- 174
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 125 GDNIRELOQRLEPYADQLRTQVNTQAEOLRQDPLAQRERERVLRENADSLQASLRPHAD 184
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 175 GLOSRYVHHHTGRFKELEHPYAESLVSGIGHVOELHRSVAPHPASPARLSRCVOYLSRK 234
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 185 ELAKKIQONVEELKGRITPYADEFKVKIIDQVEELKRSLSAPYADYDOKLHNQHLGCLTFQ 244
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 235 LTLKAKALHARIQONLDQRLRELSRAFAQT-----GTEEGAGDPQMLSEEVNQRLOAFR 289
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 245 MKKNAEELKARISASAEELRQRLAPLAEDVRGNLKGTEGLOKSLAELGHLDOQVEEFR 304
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 290 QDTYIQLIAATRAIDQTEVEYQOOLAP---PPGHSAP 324
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 305 RVEPYGENFNKALVQDMQEQIKQIKPHAGDVGHLSF 342
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||

```

RESULT 10

```

US-09-987-107-33
: Sequence 33; Application US/09987107
: Patent No. US20020156007A1
: GENERAL INFORMATION:
: APPLICANT: GRAVERSEN, Jonas
: APPLICANT: MOESTRUP, Soren
: TITLE OF INVENTION: APOLIPROTEINS ANALOGUES
: FILE REFERENCE: GRAVERSENIA
: CURRENT APPLICATION NUMBER: US/09/987,107
: CURRENT FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/264,022
: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: DK PA2001 00057
: PRIOR FILING DATE: 2001-01-15
: PRIOR APPLICATION NUMBER: DK PA2000 01682
: PRIOR FILING DATE: 2000-11-10
: NUMBER OF SEQ ID NOS: 91
: SOFTWARE: Patent In version 3.1
: SEQ ID NO 33
: LENGTH: 396
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-987-107-33

```

```

Query Match      17.4%; Score 327; DB 9; Length 396;

```

```

Best Local Similarity 28.4%; Pred. No. 5.3e-16;
Matches 96; Conservative 71; Mismatches 153; Indels 18; Gaps 6;

```

```

QY 5 AAVLTALALLSA-----FSATQARKGFMDYFSQTSQD-KGRVEDIHQOKMARE-PATLKD 58
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 5 AAVLTALVAAGARAEVSADQATVVMYDFSOI.SNNKAEVHLQKSEI.TQOLNALPHOD 64
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 59 SLEQDNLNMNKKFLEKLRPLSGSEAPRLPDPVGMRRQLOEELFEVKARLQPYMAEAHEL 118
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 65 KLGSEVNTYAGDLOKLVFPFTELHERLAKNSEKLEIGKELEELRRLPLPHANEVSQKI 124
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 119 GWNLEGIROOLKPYTMDLMEQVALRVQEIQYQOLRVGCEDTKAQILGCVDAAMALI.Q---- 174
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 125 GDNIRELOQRLEPYADQLRTQVNTQAEOLRQDPLAQRERERVLRENADSLQASLRPHAD 184
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 175 GLOSRYVHHHTGRFKELEHPYAESLVSGIGHVOELHRSVAPHPASPARLSRCVOYLSRK 234
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 185 ELAKKIQONVEELKGRITPYADEFKVKIIDQVEELKRSLSAPYADYDOKLHNQHLGCLTFQ 244
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 235 LTLKAKALHARIQONLDQRLRELSRAFAQT-----GTEEGAGDPQMLSEEVNQRLOAFR 289
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 245 MKKNAEELKARISASAEELRQRLAPLAEDVRGNLKGTEGLOKSLAELGHLDOQVEEFR 304
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 290 QDTYIQLIAATRAIDQTEVEYQOOLAP---PPGHSAP 324
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 305 RVEPYGENFNKALVQDMQEQIKQIKPHAGDVGHLSF 342
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||

```

RESULT 11

```

US-09-987-107-36
: Sequence 36; Application US/09987107
: Patent No. US20020156007A1
: GENERAL INFORMATION:
: APPLICANT: GRAVERSEN, Jonas
: APPLICANT: MOESTRUP, Soren
: TITLE OF INVENTION: APOLIPROTEINS ANALOGUES
: FILE REFERENCE: GRAVERSENIA
: CURRENT APPLICATION NUMBER: US/09/987,107
: CURRENT FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/264,022
: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: DK PA2001 00057
: PRIOR FILING DATE: 2001-01-15
: PRIOR APPLICATION NUMBER: DK PA2000 01682
: PRIOR FILING DATE: 2000-11-10
: NUMBER OF SEQ ID NOS: 91
: SOFTWARE: Patent In version 3.1
: SEQ ID NO 36
: LENGTH: 401
: TYPE: PRT
: ORGANISM: Papio anubis
US-09-987-107-36

```

```

Query Match      16.8%; Score 315; DB 9; Length 401;
Best Local Similarity 28.1%; Pred. No. 3.7e-15;
Matches 92; Conservative 70; Mismatches 137; Indels 28; Gaps 8;

```

```

QY 19 SATQARKGFMDYFSQTSQD-KGRVEDIHQOKMARE-PATLKDSELDLNMNKKFLEKLR 76
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 7 SADQVATVMMDYFSQI.SNNKAEVHLQKSEI.TQOLNALPHODKLVNRYAGDLOKRLP 66
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 77 LSGSEAPRLPDPVGMRRQLOEELFEVKARLQPYMAEAHELQMNLEGLRQOLKPYTMDL 136
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 67 FATELHERLAKDSKKLEELRKELEERARLLPHANEVSQKIENRELOQRLEPTDOL 126
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 137 MEQVALRVQEIQYQOLRVGCEDTKAQILGCVDAAMALI.Q-----GLOSRYVHHHTGRFKELE 192
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 127 RTQVNTQTEQLRQILTPYADRMERVLRENADSLQTSLRPHADOLKAKIDQNVLELGRILT 186
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
QY 193 PYAESLVSGIGHVOELHRSVAPHPASPARLSRCVOYLSRKILTLKAKALHARIQONLD 252
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 187 PYADEFKVKIIDQVEELKRSLSAPYADYDOKLHNQHLGCLTFQMKKNAEELKARISASAE 246
   ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||

```

OY 253 LREE--LSRAFACT--GTEGAGPDPOMLS-----EEVORLQARODTYLQIAAFT 300
 DB 247 LKORLAPLAEDMRGRLKNTCELOKSLAELOCHLDNRHVEERLKRVEPGEN-----FN 299
 OY 301 RAIDQTEEVQOOLAP---PPGHSAP 324
 DB 300 KALVQOMEQLRQKLGPHAGDVEGHLSF 326

RESULT 12

US-09-987-107-38
 ; Sequence 38, Application US/09987107
 ; Patent No. US20020156007A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRAVERSEN, Jonas
 ; APPLICANT: MOESTRUP, Soren
 ; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
 ; FILE REFERENCE: GRAVERSENIA
 ; CURRENT APPLICATION NUMBER: US/09/987,107
 ; CURRENT FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/264,022
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: DK PA2001 00057
 ; PRIOR FILING DATE: 2001-01-15
 ; PRIOR APPLICATION NUMBER: DK PA2000 01682
 ; PRIOR FILING DATE: 2000-11-10
 ; NUMBER OF SEQ ID NOS: 91
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 38
 ; LENGTH: 391
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 ; US-09-987-107-38

Query Match 16.6%; Score 311.5; DB 9; Length 391;
 Best Local Similarity 24.8%; Pred. No. 6.4e-15;
 Matches 95; Conservative 84; Mismatches 151; Indels 53; Gaps 9;

OY 5 AAVLTWALALISAFA--TQARKGFWDYFSQTSQD-KGRVQIHOQKMAREPATLKDS 59
 DB 5 AAVLTWALALISAFA--TQARKGFWDYFSQTSQD-KGRVQIHOQKMAREPATLKDS 61
 OY 60 LEODLNMMKFE---KLRLPSSEAPRLPDQPYGMRQLOELEEYKARLOPYMAEAH 115
 DB 62 PQDKLGNITFYADDLQNKLVPRAYQISCHLTKETERVREIQELEDLANNMHPHANKS 121
 OY 116 ELVGNMLRQLOKATYTYMDLMHOVALRVQHIQEQIRVVGCHTKAQLIGCVIPAMALLQG 175
 DB 122 QMGEDVWOKLOEHLREYATDLOAIQAOTQDMKROLTPYIQKQTTIDQNV-----EN 174
 OY 176 LOSRVVHHTGRKELPH-----PYARSIVSGICRHOEILKRSVAHPHAPASPARI 224
 DB 175 LOSRVVHHTGRKELPH-----PYARSIVSGICRHOEILKRSVAHPHAPASPARI 224
 OY 225 SRVQVYLSRKLTLKAKALHARIQOINLQLELSRAFACT--GTEGAGPDPOMLSE 279
 DB 235 NHQMEELAFQOMKKNABELQTKVSTINIDQLOKNLAPLVEDVQSKLGNTRICLOKSLIEDLAK 294
 OY 280 EVNQRQARQODTYLQIAAFTALIDQETEVQOOLAPPPGHSAPAPPEQOQDSGVY--- 336
 DB 295 QLDQOVVERRAVEPLGDFNNALVQOMKEFRQQLG-----SDSGGVESH 339
 OY 337 LSKLQARLDLMDLDTHTSLHDCG 359
 DB 340 LSLLEKNLREKVSFFMSTLQKKG 362

RESULT 13
 US-09-800-729-208
 ; Sequence 208, Application US/09800729
 ; Patent No. US20020068319A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NI et al.

; TITLE OF INVENTION: 32 Human secreted proteins
 ; FILE REFERENCE: P2044PI
 ; CURRENT APPLICATION NUMBER: US/09/800,729
 ; CURRENT FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: PCT/US00/26013
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: 60/155,709
 ; PRIOR FILING DATE: 1999-09-24
 ; NUMBER OF SEQ ID NOS: 217
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 208
 ; LENGTH: 391
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-800-729-208

Query Match 15.9%; Score 298.5; DB 10; Length 391;
 Best Local Similarity 24.4%; Pred. No. 5.2e-14;
 Matches 91; Conservative 85; Mismatches 164; Indels 33; Gaps 7;

OY 5 AAVLTWALALISAFA--FSATQARKGFWDYFSQTSQD-KGRVQIHOQKMAREPATLK-D 58
 DB 5 AAVLTWALALISAFA--FSATQARKGFWDYFSQTSQD-KGRVQIHOQKMAREPATLK-D 64
 OY 59 SLEODLNMMKFEKLRLPSSEAPRLPDQPYGMRQLOELEEYKARLOPYMAEAHLY 118
 DB 65 KLGDASTYADGVHNLVPRAYQISCHLTKETERVREIQELEDLANNMHPHANKVOTF 124
 OY 119 GWNLEGLRQLOKPYTYMDLMHOVALRVQHIQEQIRVVGCHTKAQLIGCVIPAMALLQG 174
 DB 125 GEMOKLOEHLREYATDLOAIQAOTQDMKROLTPYIQKQTTIDQNV-----EN 174
 OY 175 GLOSRVHHTGRKELPH-----PYARSIVSGICRHOEILKRSVAHPHAPASPARI 224
 DB 185 NKLDFNRMMEELKQLTPRANLTKATIDQNLQLELSRAFACT--GTEGAGPDPOMLSE 279
 OY 235 LTLKAKALHARIQOINLQLELSRAFACT--GTEGAGPDPOMLSE 279
 DB 245 MKKNABEELQTKVSAKIDQLOKNLAPLVEDVQSKLGNTRICLOKSLIEDLANNMHPHANKS 121
 OY 290 QDPTVYLSRKLTLKAKALHARIQOINLQLELSRAFACT--GTEGAGPDPOMLSE 279
 DB 305 RIVERPMGEFNKALVQOLEQFRQQLG-----NSGEVSHLSFLEKSURE 349
 OY 347 LMDLDTHTSLHDCG 359
 DB 350 KVNSTWSTLQKKG 362

RESULT 14
 US-09-987-107-35
 ; Sequence 35, Application US/09987107
 ; Patent No. US20020156007A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRAVERSEN, Jonas
 ; APPLICANT: MOESTRUP, Soren
 ; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
 ; FILE REFERENCE: GRAVERSENIA
 ; CURRENT APPLICATION NUMBER: US/09/987,107
 ; CURRENT FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/264,022
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: DK PA2001 00057
 ; PRIOR FILING DATE: 2001-01-15
 ; PRIOR APPLICATION NUMBER: DK PA2000 01682
 ; PRIOR FILING DATE: 2000-11-10
 ; NUMBER OF SEQ ID NOS: 91
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 35
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-987-107-35

Query Match 15.9%: Score 298.5; DB 9; Length 395;

Best Local Similarity 24.4%: Pred. No. 5.2e-14; Matches 91; Conservative 85; Mismatches 164; Indels 33; Gaps 7;

```

QY 5 AAVLTWALALSA-----FSATQARKGFWDYFSQTSGD-KGRVEQTHQOKMAREPATL-KD 58
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 5 AAVLTWALALSA-----FSATQARKGFWDYFSQTSGD-KGRVEQTHQOKMAREPATL-KD 64
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 59 SLFQDLNNMKNFLEKLRPLSGSEAPRLPQDPVGMRRQLOEELEEVKARIQPTMAEAHELY 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 KLGDASTYADGVHNNKLVFVYVQVLSGLHAKETERVEKEIKKELEDLRDMPIHANKVTQTF 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 GNNLEGLRQQLKPYTMDMEQVALRVQELQRLVYGEDTKQALLGYD---EAMALLQ 174
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 125 GNNMQKLOEHLKPYAVNDQDQINTQOTQKMKIQLTPYIQMOTTIKENVYDNLHTSMPLAT 184
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 175 GLQSRVHHHTGFRKELFHPYAESLVSGIRHVOELHRSVAPAPASPARLSRCVYLSRK 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 NLKDFNRRMMEELKGLHTRANELKATIDQNLDELRLSLAPLVGVQERKLNHOMEGLAFO 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 LTLKAKALHARIOQNLDOLELSRAFAGT-----GTEGAGPDPMI.SKEVROHLOAFR 289
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 245 MKKNAEELQTKYSAKIDQLOKTLAPVEDVQSKVKGNTBGLQSLLEDLRQLEQVVEFR 304
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 290 QDTYLIQIAFTRAIDQETEEVQOQLAPPPGHSAPAFPEFQOTDSKV---LSKIQARLDD 346
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 305 RTVEPMGMPFNKALVQOLEQFRQQLGP-----NSGEVESHLSFILEKSLRE 349
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 347 LMEDITHSLHDOG 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 350 KVSFPMSTLEKKG 362
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

US-09-987-107-15

Sequence 15, Application US/09987107

Patent No. US20020156007A1

GENERAL INFORMATION:

APPLICANT: GRAVERSEN, Jonas

APPLICANT: MOESTRUP, Soren

TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES

FILE REFERENCE: GRAVERSENIA

CURRENT APPLICATION NUMBER: US/09/987,107

PRIOR FILING DATE: 2001-11-13

PRIOR APPLICATION NUMBER: US 60/264,022

PRIOR FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: DK PA2001 00057

PRIOR FILING DATE: 2001-01-15

PRIOR APPLICATION NUMBER: DK PA2000 01682

PRIOR FILING DATE: 2000-11-10

NUMBER OF SEQ ID NOS: 91

SOFTWARE: PatentIn version 3.1

SEQ ID NO 15

LENGTH: 267

TYPE: PRT

ORGANISM: Homo sapiens

US-09-987-107-15

Query Match 10.7%: Score 200; DB 9; Length 267;

-6 30 14 258" data-label="Text">

Best Local Similarity 22.2%: Pred. No. 2.6e-07;

Matches 74; Conservative 61; Mismatches 106; Indels 92; Gaps 10;

```

QY 5 AAVLTWALALSA-----FSATQARKGFWDYFSQTSGDGR--VEQTH 45
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 3 AAVLTWALALSA-----FSATQARKGFWDYFSQTSGDGR--VEQTH 58
   ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
QY 46 QOKMARE-PATLKDSLQEDLNMMNFKELRPLSGSEAPRLPQDPVGMRRQLOEELEEVK 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 GSAICGKQNLKLLDMWDSVTSTFSKLRQGLGVYQGFMDNLEKTEGRLRQEMSKDLEEVK 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 105 ARLOPYMAEAHELYGNNLEGLRQQLKPYTMDMEQVALRVQELQRLVYGEDTKQALLG 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 119 AKVQPYLDDFOKKMOEMELRYRQKVEPLRAELQEGARQKHLHELQKSLPLGEEMR----- 173
QY 165 GVDAMALLQGLQSRVHHHTGFRKELFHPYAESLVSGIRHVOELHRSVAPAPASPARL 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 --DKARA-----HVDALRTHLAPYSDLRQL 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 SRCVQVLSRKLTLKAKALHARIOQNLDOLELSRAFAGTGTEGAGPDPMI.SKEVROH 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 199 AARLEALKENGARLAELYHAKATEHLSTLSEKAKPAL-----EDLRQK 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 285 ---LQAFRQDTYLIQIAFTRAIDQETEEVQO 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 LLPVLESEK-----VSFLSALLEETKLNQ 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: January 23, 2003, 17:48:14

Job time : 41 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2003, 17:41:21 : Search time 38 Seconds

(without alignments)
925.926 Million cell updates/sec

Title: US-09-842-364-3

Perfect score: 1877

Sequence: 1 MASMAVLTWMLALLSASFSA.....LMEDITHSLHDGSHLGD 366

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : PIR-73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	333	17.7	429	2	S29565
2	328	17.5	396	1	IAPHUA4
3	319.5	17.0	391	1	IAPRTA4
4	315	16.8	401	2	A47141
5	298.5	15.9	391	2	B40892
6	298.5	15.9	395	2	A40892
7	289.5	15.4	399	2	C40892
8	288	15.3	394	2	A25281
9	200	10.7	267	1	IAPHUA1
10	193	10.3	267	1	JS0079
11	187	10.0	267	1	A26529
12	185	9.9	266	1	LPRB12
13	183	9.7	266	1	LPRB12
14	180	9.6	265	1	LPRB18
15	178	9.5	265	2	AC5858
16	177	9.4	329	2	AC5856
17	175.5	9.4	266	1	LPGD11
18	170.5	9.1	241	2	A24998
19	170	9.1	265	2	JT0672
20	168	9.0	265	2	A46018
21	168	9.0	317	2	A28792
22	168	9.0	317	2	S03185
23	165	8.8	317	2	S33450
24	163.5	8.7	1162	2	D83454
25	161	8.6	264	2	S22420
26	158.5	8.4	246	2	A61448
27	158.5	8.4	264	2	S31394
28	155	8.3	264	2	UC5456
29	155	8.3	311	2	JU0036

30	154.5	8.2	262	2	JC1237	apolipoprotein A-I
31	154.5	8.2	317	1	LPRRE	apolipoprotein E p
32	153	8.2	317	1	LPHOE	apolipoprotein E p
33	152.5	8.1	291	2	C60940	apolipoprotein E p
34	152.5	8.1	298	2	S12635	apolipoprotein E p
35	152.5	8.1	986	2	T10754	cis-Golgi matrix p
36	151.5	8.1	1956	2	T16416	hypothetical prote
37	151	8.0	259	2	A24700	apolipoprotein A-I
38	150.5	8.0	316	2	S26478	apolipoprotein E -
39	150.5	8.0	1034	2	T32297	hypothetical prote
40	149	7.9	311	2	A45951	apolipoprotein E p
41	147.5	7.9	231	2	J00704	apolipoprotein A-I
42	147.5	7.9	316	2	I45996	apolipoprotein E -
43	147.5	7.9	1248	2	AH2637	conserved hypothet
44	146	7.8	316	2	JC6549	apolipoprotein E p
45	144.5	7.7	164	2	S21830	apolipoprotein A-I

ALIGNMENTS

RESULT 1

S29565

apolipoprotein A-IV - crab-eating macaque

C:Species: Macaca fascicularis (crab-eating macaque)

C:Date: 06-Jan-1995 #sequence-revision 06-Jan-1995 #text-change 13-Aug-1999

C:Accession: S30195; S29565

R:Osada, J., Pocovi, M., Nicolson, R.J., Schaefer, E.J., Ordovas, J.M.

Biochim. Biophys. Acta 1172, 335-339, 1993

A:Title: Nucleotide sequences of the Macaca fascicularis apolipoprotein C-III and A-

A:Reference number: S30195; MID:93192330; PMID:8448212

A:Accession: S30195

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1429 <OSAS>

A:Cross-references: EMBL:X68361; NID:q38050; PIDN:CAA48421.1; PID:q38051

C:Genetics:

A:Insertions: 17/1: 59/2

C:Superfamily: apolipoprotein A-I

Query Match 17.7%; Score 333; DB 2; Length 429;

Best Local Similarity 28.7%; Pred. No. 4.5e-13;

Matches 99; Conservative 72; Mismatches 142; Indels 32; Gaps 9;

QY	5	AAVLTWMLALISA---FSATQARKGFWDYFSQTSQD-KGRVEDIHOQKARE-PATLKD	58
DB	5	AAVLTWMLALISA---FSATQARKGFWDYFSQTSQD-KGRVEDIHOQKARE-PATLKD	58
QY	59	SLEODLNMMKFLKRLPLSGSEAPRLPOPVGMRROLOELEVYKARLOPYMAEAEIV	118
DB	59	SLEODLNMMKFLKRLPLSGSEAPRLPOPVGMRROLOELEVYKARLOPYMAEAEIV	118
QY	65	KIDVNTYVAGDLQKLVFPFATELHEKLAKEIKLELEVERKRLPHANEVSOKI	124
DB	65	KIDVNTYVAGDLQKLVFPFATELHEKLAKEIKLELEVERKRLPHANEVSOKI	124
QY	119	GMNLEGLRQOLKPYTMDLMEQVALRVOELQOLRVVEEDPKAOLGVDPMALIQ----	174
DB	119	GMNLEGLRQOLKPYTMDLMEQVALRVOELQOLRVVEEDPKAOLGVDPMALIQ----	174
QY	125	GENVRELDQGLLEPTDQLRQVNTQTEQLRQGLTRYQNRNERYLRNADLSLRHAD	184
DB	125	GENVRELDQGLLEPTDQLRQVNTQTEQLRQGLTRYQNRNERYLRNADLSLRHAD	184
QY	175	GIOSRVVNHGTRKELFHPYAESLVSGTGRHVOELHRSVAPHPAPASARLSKVOVLSRK	234
DB	175	GIOSRVVNHGTRKELFHPYAESLVSGTGRHVOELHRSVAPHPAPASARLSKVOVLSRK	234
QY	185	QLAKAKIDQNEVELKERLTLPYADEFKVYIDQTEVELRSLSLAPYQDDAEKLNHLEGLAFQ	244
DB	185	QLAKAKIDQNEVELKERLTLPYADEFKVYIDQTEVELRSLSLAPYQDDAEKLNHLEGLAFQ	244
QY	235	LTGKAKALHARIOONDLQREE---LSRAVAGT--GTGEGGDPQQLS-----EVR	282
DB	235	LTGKAKALHARIOONDLQREE---LSRAVAGT--GTGEGGDPQQLS-----EVR	282
QY	245	MKNNAEELKARISASAEELQRLAPLAEDMKCNLRGTEGLQSLAEFLGCHLDNRHVEFR	304
DB	245	MKNNAEELKARISASAEELQRLAPLAEDMKCNLRGTEGLQSLAEFLGCHLDNRHVEFR	304
QY	283	QRQDAFRQDTYLOIAETRAIDQTEFEVQQLAP--PPGHSNF	324
DB	283	QRQDAFRQDTYLOIAETRAIDQTEFEVQQLAP--PPGHSNF	324
QY	305	LREVEYGEN-----FNKALVOOMEQLRQKLPFHAGDVGHTLSF	342
DB	305	LREVEYGEN-----FNKALVOOMEQLRQKLPFHAGDVGHTLSF	342

RESULT 2

LPHUA4

apolipoprotein A-IV precursor [validated] - human

N:Alternate names: apoA-IV

C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1997 #sequence, revision 30-Sep-1997 #text, change 08-Dec-2000
 C:Accession: A94137; A94059; A24449; A29330; A26280; 137177; C54223; A61203; A26481; S02
 R:Karathanasis, S.K.; Oetgen, P.; Haddad, I.A.; Antonarakis, S.E.
 A:Title: Acad. Sci. U.S.A. 83, 8437-8461, 1986
 A:Title: Structure, evolution, and polymorphisms of the human apolipoprotein A4 gene (AP
 A:Reference number: A94137; MUID:87041474; PMID:3095836
 A:Accession: A94137
 A:Molecule type: DNA
 A:Residues: 1-396 <KAR1>
 A:Cross-references: GB:M4642; NID:q178760; PIDN:AAA51745.1; PID:q178761
 R:Karathanasis, S.K.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6374-6378, 1985
 A:Title: Apolipoprotein multigene family: tandem organization of human apolipoprotein A4
 A:Reference number: A94059; MUID:86016704; PMID:3931073
 A:Accession: A94059
 A:Molecule type: mRNA
 A:Residues: 135-378 <KAR2>
 A:Cross-references: GB:M10373; NID:q563319; PIDN:AA859516.1; PID:q563320
 R:Karathanasis, S.K.; Yunis, J.
 Biochemistry 25, 3962-3970, 1986
 A:Title: Structure, evolution, and tissue-specific synthesis of human apolipoprotein AIV
 A:Reference number: A24449; MUID:86296629; PMID:3755616
 A:Accession: A24449
 A:Molecule type: mRNA
 A:Residues: 1-396 <KAR3>
 A:Cross-references: GB:M13654; NID:q178758; PIDN:AAA51744.1; PID:q178759
 R:Gordon, J.I.; Bisgaler, C.L.; Sims, H.F.; Sachdev, O.P.; Glickman, R.M.; Strauss, A.W.
 J. Biol. Chem. 259, 468-474, 1984
 A:Title: Biosynthesis of human preapoprotein A-IV
 A:Reference number: A92475; MUID:84161950; PMID:6706947
 A:Contents: annotation: signal sequence cleavage site
 R:Elshourbagy, N.A.; Walker, D.W.; Paik, Y.K.; Boguski, M.S.; Freeman, M.; Gordon, J.I.;
 J. Biol. Chem. 262, 7973-7981, 1987
 A:Title: Structure and expression of the human apolipoprotein A-IV gene.
 A:Reference number: A29330; MUID:87250378; PMID:3036793
 A:Accession: A29330
 A:Molecule type: DNA
 A:Residues: 1-157, 'T', 159, 'Y', 161-278, 'R', 280-326, 'T', 328-379, 'H', 381-396 <ELS>
 A:Cross-references: GB:J02758; NID:q178756; PIDN:AAA96731.1; PID:q178757
 R:Elshourbagy, N.A.; Walker, D.W.; Boguski, M.S.; Gordon, J.I.; Taylor, J.M.
 J. Biol. Chem. 261, 1998-2002, 1986
 A:Title: The nucleotide and derived amino acid sequence of human apolipoprotein A-IV mRNA
 A:Reference number: A26280; MUID:86111885; PMID:3080432
 A:Accession: A26280
 A:Molecule type: mRNA
 A:Residues: 21-157, 'T', 159, 'Y', 161-278, 'R', 280-326, 'T', 328-379, 'H', 381-396 <EL2>
 A:Cross-references: GB:M14566; NID:q178778; PIDN:AAA51748.1; PID:q178779
 R:Yang, C.Y.; Gu, Z.W.; Chong, I.S.; Xiong, W.J.; Kosseneu, M.; Yang, H.X.; Lee, B.R.; G
 Biochim. Biophys. Acta 1002, 231-237, 1989
 A:Title: The primary structure of human apolipoprotein A-IV.
 A:Reference number: I37177; MUID:89194198; PMID:2930771
 A:Accession: I37177
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-157, 'T', 159, 'Y', 161-278, 'R', 280-396 <YAN1>
 A:Cross-references: EMBL:X13629; NID:928761; PIDN:CAA31955.1; PID:q28762
 R:Note: submitted to the EMBL Data Library, January 1989
 R:Knittake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.
 Biochemistry 33, 1988-1993, 1994
 A:Title: Identification of proteins associated with the human apolipoprotein A-IV containing lipop
 A:Reference number: A54223; MUID:94162201; PMID:8117655
 A:Accession: C54223
 A:Molecule type: Protein
 A:Residues: 'X', 22, 'X', 24, 'X', 26-31, 'X', 33-34 <KUN>
 R:Tenkunen, H.; Lukka, M.; Jautilainen, M.; Metsä, J.; Baumann, M.; Peltonen, L.; Ehnholm
 Arterioscler. Thromb. 11, 851-856, 1991
 A:Title: The mutation causing the common apolipoprotein A-IV polymorphism is a glutamine
 A:Reference number: A61203; MUID:91291788; PMID:2065039
 A:Accession: A61203
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 378-379, 'H', 381-382 <REN>

C:Comment: ApoA-IV is a major lipoprotein of lymph chylomicrons. In human plasma, it
 C:Comment: ApoA-IV is synthesized primarily in the intestine.
 C:GeneID: 858
 A:Gene: GDB:AP04
 A:Cross-references: GDB:119000; OMIM:107690
 A:Map position: 11q23-11q23
 A:Introns: 17/1; 59/2
 C:Superfamily: apolipoprotein A-I
 C:Keywords: chylomicron; HDL; intestine; lipid binding; lipid transport; lipoprotein
 F:1-20/Domain: signal sequence #status experimental <SIG>
 F:21-396/Product: apolipoprotein A-IV #status experimental <MAT>

Query Match 17.5% Score 328; DB 1; Length 396;
 Best Local Similarity 28.4%; Pred. No. 8.2e-13;
 Matches 96; Conservative 72; Mismatches 152; Indels 18; Gaps 6;

QY 5 AAVLTALALLSA----FSATQARKGFMDYFSQTSQD-KGRVEQIHQKMKRE-PATLKD 58
 DB 5 AAVLTALALLVAAGARAVSADQVATVMVDYFQSLNNKAEVHLQKSELTQOLNALFQD 64
 QY 59 SLEQDNNMKNKFLKRLPLSGSEAPRLPODPVGMROELEEYKARLOPYMAEAEHLY 118
 DB 65 KLGEVNTYAGDLQKLVFAPFELHRLAKDSEKLEKELEELRALLLPHANEVSOKI 124
 QY 119 GNNLEGLNQOLKPYTMDLMEQVALRVEQLQQLRVYGEQTAQLLGVDEAMALLQ--- 174
 DB 125 GDNLEELQOORLEPYADQRLTQVNTQAEQLRQDLPLAORMEVLRENADSLQASIRPHAD 184
 QY 175 GLOSIVVHTGREFKELFHPYAESLVSGIGRIYVQELHRSVAPAPASPARLSRCVOYLSRK 234
 DB 185 ELKATIDQNEVELKQRLTPYADERKVKIDQVTELRSLAPADQTKLWHQLETPQ 244
 QY 235 LTKAKALHARIQNLQDLREELSRFAQT-----GTEEGAGPPQMLSEEVORQLQAFR 289
 DB 245 MKKNEELKARISASEELRQRLAPLADVDGKLNKGTGLOKSLAELGHLDDQVEEER 304
 QY 290 QDTYQIALAFRAIDQETEEVQOQLAP---PPHGSAP 324
 DB 305 RVEPYGENFMKALVQOMEQLQKLGPHAGVEGLHSF 342

RESULT 3

LIPRA4

apolipoprotein A-IV precursor - rat

C:Species: Rattus norvegicus (norway rat)

C:Date: 27-Nov-1985 #sequence, revision 27-Nov-1985 #text, change 22-Jun-1999

C:Accession: A03095; A25214; C24700

R:Boguski, M.S.; Elshourbagy, N.; Taylor, J.M.; Gordon, J.I.

Proc. Natl. Acad. Sci. U.S.A. 81, 5021-5025, 1984

A:Title: Rat apolipoprotein A-IV contains 13 tandem repetitions of a 22-amino acid s

A:Reference number: A03095; MUID:84298074; PMID:6591177

A:Accession: A03095

A:Molecule type: mRNA

A:Residues: 1-391 <BOG>

A:Cross-references: GB:M00002; GB:K02421; NID:q202949; PIDN:AAA5909.1; PID:q202950

R:Boguski, M.S.; Birkenmeier, E.H.; Elshourbagy, N.A.; Taylor, J.M.; Gordon, J.I.

J. Biol. Chem. 261, 6398-6407, 1986

A:Title: Evolution of the apolipoproteins. Structure of the rat Apo-A-IV gene and its

A:Reference number: A25214; MUID:86196059; PMID:3009456

A:Accession: A25214

A:Molecule type: Protein

A:Residues: 1-252, 'Q', 254-391 <BOZ>

R:Haddad, I.A.; Ordovas, J.M.; Fitzpatrick, T.; Karathanasis, S.K.

J. Biol. Chem. 261, 13268-13277, 1986

A:Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and

A:Reference number: A92558; MUID:87008540; PMID:3020028

A:Accession: C24700

A:Molecule type: DNA

A:Residues: 1-252, 'Q', 254-391 <HAD>

A:Cross-references: GB:J02588; NID:q202937; PIDN:AAA40747.1; PID:q202941

C:Comment: This apoprotein is a major component of HDL and chylomicrons but, unlike c

C:Comment: Nine of the thirteen 22-amino acid tandem repeats (each 22-mer is actually

cal, and many of these helices are amphipathic. They may therefore serve as lipid-bit

C:Superfamily: apolipoprotein A-I
 C:Keywords: Chylomicron; duplication; HDL; lipid transport; plasma; tandem repeat
 F:1-20/Domains: signal sequence #status predicted <SIG>
 F:21-391/Product: apolipoprotein A-IV #status predicted <MAT>
 F:33-330/Region: 22-residue repeats

Query Match 17.0%; Score 319.5; DB 1; Length 391;
 Best Local Similarity 25.1%; Pred. No. 2,66-12;
 Matches 96; Conservative 84; Mismatches 150; Indels 53; Gaps 9;

5 AAVLTWALALSAFA---TQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLKUS 59
 DB AAVLTWALALSAFA---TQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLKUS 59
 5 AAVLTWALALSAFA---TQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLKUS 59
 DB AAVLTWALALSAFA---TQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLKUS 59
 60 LEQDNNMKKFLF---KLRPLSGSEAPRLPDVGMRRQIQLEELFVVARLOPYMAEAH 115
 DB LEQDNNMKKFLF---KLRPLSGSEAPRLPDVGMRRQIQLEELFVVARLOPYMAEAH 115
 62 FQDKLCINITYADUIONKLVPPAVOLSGHLTETEKREKEIOKELELIKANNMPHANKVS 121
 DB FQDKLCINITYADUIONKLVPPAVOLSGHLTETEKREKEIOKELELIKANNMPHANKVS 121
 116 ELVGNLEGLRQOLKPYTDLMEQVALRVQLEQLELVGVEDTKAQLGCVDEAMALLQG 175
 DB ELVGNLEGLRQOLKPYTDLMEQVALRVQLEQLELVGVEDTKAQLGCVDEAMALLQG 175
 122 QMGDNVQKLEQELRLRYAVDLQAIINAGQDMKRQQLTPYIQHMQTTLQUNV-----EN 174
 DB QMGDNVQKLEQELRLRYAVDLQAIINAGQDMKRQQLTPYIQHMQTTLQUNV-----EN 174
 176 LOSRVYHNGRKELPH-----PYAESVSGICRHVQELHRSVAPHPASAPAKI 224
 DB LOSRVYHNGRKELPH-----PYAESVSGICRHVQELHRSVAPHPASAPAKI 224
 175 LOSRVYHNGRKELPH-----PYAESVSGICRHVQELHRSVAPHPASAPAKI 234
 DB LOSRVYHNGRKELPH-----PYAESVSGICRHVQELHRSVAPHPASAPAKI 234
 225 SRGVYLSKRLTLKAKALHARIQONIDQLEELSRPAFCT-----GTEGAGHPQMLSE 279
 DB SRGVYLSKRLTLKAKALHARIQONIDQLEELSRPAFCT-----GTEGAGHPQMLSE 279
 235 NQMEELAFQMKNAEELHTKYSTNIDQIKNLAPLEVQSKIKACGTGLQKSLDELAK 294
 DB NQMEELAFQMKNAEELHTKYSTNIDQIKNLAPLEVQSKIKACGTGLQKSLDELAK 294
 280 EVRQRLQAFRQDTYQIQAFTRAIDQETEEVQOQLAPPPGHSAPAFEPQDTSKGV--- 336
 DB EVRQRLQAFRQDTYQIQAFTRAIDQETEEVQOQLAPPPGHSAPAFEPQDTSKGV--- 336
 295 QLDQGVFVRRAVEPLGDFNMALVQOMKFRQOLG-----SDSGGVESH 339
 DB QLDQGVFVRRAVEPLGDFNMALVQOMKFRQOLG-----SDSGGVESH 339
 337 LSKLARLDLMDLMDITHSLHDOG 359
 DB LSKLARLDLMDLMDITHSLHDOG 359
 340 LSTLEKLEKESVFSMTLQKKG 362
 DB LSTLEKLEKESVFSMTLQKKG 362

RESULT 4

apolipoprotein A-IV I isoform - baboon (fragment)
 C:Species: Papio sp. (baboon)
 C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: A47141
 R:Hiexon, J.E.; Kammerer, C.M.; Mott, G.E.; Britten, M.J.; Birnbaum, S.; Powers, P.K.; V.
 J. Biol. Chem. 268, 15667-15673, 1993
 A:Title: Baboon apolipoprotein A-IV. Identification of Lys76-->Glu that distinguishes the
 A:Reference number: A47141; MUID:93340170; PMID:8101842
 A:Accession: A47141
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-401 <HIK>
 A:Experimental source: intestine
 A:Note: sequence extracted from NCBI backbone (NCBIN:136009, NCBIPI:136010)
 C:Superfamily: apolipoprotein A-I

Query Match 16.8%; Score 315; DB 2; Length 401;
 Best Local Similarity 28.1%; Pred. No. 5,1e-12;
 Matches 92; Conservative 70; Mismatches 137; Indels 28; Gaps 8;

19 SATQARKGFWDYFSQSGD-KGRVEQIHQOKMARE-PATLKDSLQEDLNMMNKFLKLRP 76
 DB SATQARKGFWDYFSQSGD-KGRVEQIHQOKMARE-PATLKDSLQEDLNMMNKFLKLRP 76
 7 SADVATVWMDYFSQSSNAKAEVHLOKSELQQLNALPQDKLGEVNTYAGLOKKIAP 66
 DB SADVATVWMDYFSQSSNAKAEVHLOKSELQQLNALPQDKLGEVNTYAGLOKKIAP 66
 77 LSGSEAPRLPDVGMRRQIQLEELFVVARLOPYMAEAHVGWNLGLRQOLKPYTMDI 136
 DB LSGSEAPRLPDVGMRRQIQLEELFVVARLOPYMAEAHVGWNLGLRQOLKPYTMDI 136
 67 FATELHERLAKQSKKLEELRKELFVARLPLHANEVQKIGENRELOQRLEPYTDOL 126
 DB FATELHERLAKQSKKLEELRKELFVARLPLHANEVQKIGENRELOQRLEPYTDOL 126
 137 MEQVALRVQLEQLELVGVEDTKAQLGCVDEAMALLQ-----GLOSRYVHNGRKELPH 192
 DB MEQVALRVQLEQLELVGVEDTKAQLGCVDEAMALLQ-----GLOSRYVHNGRKELPH 192
 127 KRVQNTQTEOLRQQLTPYIQHMQTTLQUNV-----EN 186
 DB KRVQNTQTEOLRQQLTPYIQHMQTTLQUNV-----EN 186

193 PYAESVSGICRHVQELHRSVAPHPASAPRLSGVOYLSRKLTLKAKALHARIQONIDQ 252
 DB PYAESVSGICRHVQELHRSVAPHPASAPRLSGVOYLSRKLTLKAKALHARIQONIDQ 252
 187 PVADEFKKTIDQTEVELRSLAPYADQAEKILNQLEGLAFQMKNAEELKARISASAP 246
 DB PVADEFKKTIDQTEVELRSLAPYADQAEKILNQLEGLAFQMKNAEELKARISASAP 246
 253 LKEE---LSRAPACT---GTEGAGHPQMLSE-----EVRQRLQAFRQDTYQIQAFT 300
 DB LKEE---LSRAPACT---GTEGAGHPQMLSE-----EVRQRLQAFRQDTYQIQAFT 300
 247 LQORLAPLAEDRGNRLRGTEGLQSLAELGILDRHVEEFLRVLEPYGEN-----FN 299
 DB LQORLAPLAEDRGNRLRGTEGLQSLAELGILDRHVEEFLRVLEPYGEN-----FN 299
 301 KALIDETEVOOQIAP---PPGHSAP 324
 DB KALIDETEVOOQIAP---PPGHSAP 324
 300 KALVQOMKFRQOLGDFNMALVQOMKFRQOLG-----SDSGGVESH 326
 DB KALVQOMKFRQOLGDFNMALVQOMKFRQOLG-----SDSGGVESH 326

RESULT 5

apolipoprotein A-IV precursor - mouse (strain 129)
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Mar-1992 #sequence_revision 10-Apr-1992 #text_change 13-Aug-1999
 C:Accession: B40892
 R:Reue, K.; Leete, T.H.
 J. Biol. Chem. 266, 12715-12721, 1991
 A:Title: Genetic variation in mouse apolipoprotein A-IV due to insertion and deletion
 A:Reference number: B40892; MUID:91286309; PMID:1648102
 A:Accession: B40892
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-391 <REU>
 A:Cross-references: GB:M4248; NID:919184; PID:NAA37214.1; PID:919185
 A:Note: the authors translated the codon CTG for residue 87 as Glu, CAG for residue
 C:Superfamily: apolipoprotein A-I
 C:Keywords: lipid binding; lipoprotein

Query Match 15.9%; Score 298.5; DB 2; Length 391;
 Best Local Similarity 24.4%; Pred. No. 4,9e-11;
 Matches 91; Conservative 85; Mismatches 164; Indels 33; Gaps 7;

5 AAVLTWALALSAFA---FSATQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLK-D 58
 DB AAVLTWALALSAFA---FSATQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLK-D 58
 5 AAVLTWALALSAFA---FSATQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLK-D 64
 DB AAVLTWALALSAFA---FSATQARKGFWDYFSQSGD-KGRVEQIHQOKMAREPATLK-D 64
 59 SLKQDNNMKKFLF---KLRPLSGSEAPRLPDVGMRRQIQLEELFVVARLOPYMAEAH 118
 DB SLKQDNNMKKFLF---KLRPLSGSEAPRLPDVGMRRQIQLEELFVVARLOPYMAEAH 118
 65 KLGDASTYADVGNKLVPPAVOLSGHLTETEKREKEIOKELELIKANNMPHANKVS 124
 DB KLGDASTYADVGNKLVPPAVOLSGHLTETEKREKEIOKELELIKANNMPHANKVS 124
 119 GWNLEGLRQOLKPYTDLMEQVALRVQLEQLELVGVEDTKAQLGCVDEAMALLQ 174
 DB GWNLEGLRQOLKPYTDLMEQVALRVQLEQLELVGVEDTKAQLGCVDEAMALLQ 174
 125 GRNMOKIOELHRIKRYAVDLQAIINAGQDMKRQQLTPYIQHMQTTLQUNV-----EN 184
 DB GRNMOKIOELHRIKRYAVDLQAIINAGQDMKRQQLTPYIQHMQTTLQUNV-----EN 184
 175 GLOSRYVHNGRKELPH-----PYAESVSGICRHVQELHRSVAPHPASAPRLSGVOYLSR 234
 DB GLOSRYVHNGRKELPH-----PYAESVSGICRHVQELHRSVAPHPASAPRLSGVOYLSR 234
 185 NIKDKFNKMMKELKGLTTRANLKAFTIIONLELIKRSIAPLTVGQELNMQEGLAFQ 244
 DB NIKDKFNKMMKELKGLTTRANLKAFTIIONLELIKRSIAPLTVGQELNMQEGLAFQ 244
 235 LTLKAKALHARIQONIDQLEELSRPAFCT-----GTEGAGHPQMLSEVRQRLQAF 289
 DB LTLKAKALHARIQONIDQLEELSRPAFCT-----GTEGAGHPQMLSEVRQRLQAF 289
 245 MKKNAEELQTKVSADQLOKKNLAPLEVQSKIKACGTGLQKSLDELAK 304
 DB MKKNAEELQTKVSADQLOKKNLAPLEVQSKIKACGTGLQKSLDELAK 304
 290 QDTYQIQAFTRAIDQETEEVQOQLAPPPGHSAPAFEPQDTSKGV---LSKLQARLD 346
 DB QDTYQIQAFTRAIDQETEEVQOQLAPPPGHSAPAFEPQDTSKGV---LSKLQARLD 346
 305 RIVPEGEMFNALVQOMKFRQOLG-----SDSGGVESH 326
 DB RIVPEGEMFNALVQOMKFRQOLG-----SDSGGVESH 326
 347 LMDITHSLHDOG 359
 DB LMDITHSLHDOG 359
 350 KVSFMSSTLEKKG 362
 DB KVSFMSSTLEKKG 362

RESULT 6

apolipoprotein A-IV precursor - mouse (strain C57BL/6)
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Mar-1992 #sequence_revision 10-Apr-1992 #text_change 13-Aug-1999
 C:Accession: A40892
 R:Reue, K.; Leete, T.H.
 J. Biol. Chem. 266, 12715-12721, 1991

RESULT 9

LEPRAI

apolipoprotein A-I precursor [validated] - human

N:Alternative names: APOA-I-2; APOA-I-4; preproapoA-I; proslacylin stabilizing factor

C:Species: Homo sapiens (man)

C:Date: 14-Nov-1983 #sequence, revision 14-Nov-1983 #text, change 08-Dec-2000

C:Accession: A90947; S02373; A93465; A21147; A93519; B93519; A93472; A94010; A21

R:Selkhamer, J.J.; Procter, A.A.; Frossard, P.; Levy-Wilson, B.

DNA 3, 309-317, 1984

A:Title: Isolation and DNA sequence of full-length cDNA and of the entire gene for human

A:Reference number: A90947; MUID:85026665; PMID:6207999

A:Accession: A90947

A:Molecule type: DNA

A:Residues: 1-267 <SEI>

A:Cross-references: GB:J01038; NID:q28769; PIDN:CAA25519.1; PID:q296635

A:Accession: B90947

A:Molecule type: mRNA

A:Residues: 1-267 <SE2>

A:Cross-references: GB:X01038; NID:q28769; PIDN:CAA25519.1; PID:q296635

R:Markides, S.C.; Ruiz-Opazo, N.; Hayden, M.; Nusbaum, A.L.; Breslow, J.L.; Zannis, V.I.

Eur. J. Biochem. 173, 465-471, 1988

A:Title: Sequence and expression of Tangier apoA-I gene.

A:Reference number: S02373; MUID:88196137; PMID:3129297

A:Accession: S02373

A:Molecule type: DNA

A:Residues: 1-267 <MAK>

A:Cross-references: EMBL:X07496; NID:q28774; PIDN:CAA30377.1; PID:q296729

R:Shoulders, C.C.; Kornblith, A.R.; Munro, B.S.; Battelle, F.E.

Nucleic Acids Res. 11, 2827-2837, 1983

A:Title: Gene structure of human apolipoprotein AI.

A:Reference number: A93465; MUID:83220822; PMID:6406984

A:Accession: A93465

A:Molecule type: DNA

A:Residues: 1-267 <SHO>

A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB

R:Karathanasis, S.K.; Zannis, V.I.; Breslow, J.L.

Proc. Natl. Acad. Sci. U.S.A. 80, 6147-6151, 1983

A:Title: Isolation and characterization of the human apolipoprotein A-I gene.

A:Reference number: A21147; MUID:84016011; PMID:6413973

A:Accession: A21147

A:Molecule type: DNA

A:Residues: 1-267 <KAR>

A:Cross-references: GB:J00098; GB:J03222; NID:q178765; PIDN:AA859514.1; PID:q178768

R:Shapiro, C.R.; Sidoli, A.; Shelley, C.S.; Lucero, M.A.; Shoulters, C.C.; Battelle, F.E.

Nucleic Acids Res. 12, 3917-3932, 1984

A:Title: Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA abundance.

A:Reference number: A93519; MUID:84221405; PMID:6328445

A:Accession: A93519

A:Molecule type: mRNA

A:Residues: 1-267 <SHA>

A:Cross-references: GB:X00566; NID:q28765; PIDN:CAA25332.1; PID:q372753

A:Accession: B93519

A:Molecule type: DNA

A:Residues: 1-24 <SH2>

R:Cheung, P.; Chan, L.

Nucleic Acids Res. 11, 3703-3715, 1983

A:Title: Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.

A:Reference number: A93472; MUID:83220772; PMID:6304641

A:Accession: A93472

A:Molecule type: mRNA

A:Residues: 1-267 <CHE>

A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB

R:Law, S.W.; Brewer Jr., H.B.

Proc. Natl. Acad. Sci. U.S.A. 81, 66-70, 1984

A:Title: Nucleotide sequence and the encoded amino acids of human apolipoprotein A-I mRNA

A:Reference number: A94010; MUID:84119464; PMID:6198645

A:Accession: A94010

A:Molecule type: mRNA

A:Residues: 1-267 <LAW>

A:Cross-references: GB:J00098; GB:J00099; GB:J00100; GB:J00101; GB:J03222; GB:K01518; GB

R:Zannis, V.I.; Karathanasis, S.K.; Neumann, H.T.; Goldberger, G.; Breslow, J.L.

Proc. Natl. Acad. Sci. U.S.A. 80, 2574-2578, 1983

A:Title: Intracellular and extracellular processing of human apolipoprotein A-I: secret

A:Reference number: A21118; MUID:83195100; PMID:6405383

A:Accession: A21118

A:Molecule type: mRNA

A:Residues: 1-24 <ZAN>

R:Brewer Jr., H.B.; Fairwell, T.; Kay, L.; Meng, M.; Roman, R.; Law, S.; Light, J.A.

Biochem. Biophys. Res. Commun. 113, 626-632, 1983

A:Title: Human plasma proapoA-I: isolation and amino-terminal sequence.

A:Reference number: A90112; MUID:83256553; PMID:6409108

A:Accession: A90112

A:Molecule type: protein

A:Residues: 19-27 <HRE>

R:Brewer Jr., H.B.; Fairwell, T.; Lank, A.; Roman, R.; Houser, A.; Bronzert, T.J.

Biochem. Biophys. Res. Commun. 80, 623-630, 1978

A:Title: The amino acid sequence of human APOA-I, an apolipoprotein isolated from hlg

A:Reference number: A90209; MUID:78123731; PMID:204308

A:Accession: A90209

A:Molecule type: protein

A:Residues: 25-57, 'Q', '59-169, 'QO', '172-267 <HR2>

R:Yui, Y.; Aoyama, T.; Morishita, H.; Takahashi, M.; Takatsu, Y.; Kawai, C.

J. Clin. Invest. 82, 803-807, 1988

A:Title: Serum proslacylin stabilizing factor is identical to apolipoprotein A-I (AI

A:Reference number: A30516; MUID:88331387; PMID:3047170

A:Accession: A30516

A:Molecule type: protein

A:Residues: 25-56 <YUI>

R:Nichols, W.C.; Dwyer, F.E.; Liepnies, J.; Benson, M.D.

Biochem. Biophys. Res. Commun. 156, 762-768, 1988

A:Title: Variant apolipoprotein AI as a major constituent of a human hereditary amyic

A:Reference number: A31582; MUID:89050104; PMID:3142462

A:Accession: A31582

A:Molecule type: protein

A:Residues: 25-49, 'R', '51-85, 'D', '87-107 <NIC>

R:Manjunath, P.; Marcel, Y.L.; Ume, J.; Seidah, N.G.; Chretien, P.; Chappelaine, A.

J. Biol. Chem. 264, 16853-16857, 1989

A:Title: Apolipoprotein A-I binds to a family of bovine seminal plasma proteins.

A:Reference number: A34409; MUID:89380318; PMID:2506184

A:Accession: A34409

A:Molecule type: protein

A:Residues: 25-48 <MAN>

R:Stoffel, W.; Binczek, E.

Biol. Chem. Hoppe-Seyler 369, 1055-1063, 1988

A:Title: Structural requirements of human preproapolipoprotein AI for translocation t

A:Reference number: S02377; MUID:89149957; PMID:3228490

A:Accession: S02377

A:Molecule type: protein

A:Residues: 1-6, 'AV', '9, '1V', '12-29 <STO>

A:Note: part of this sequence, including the amino end of the mature protein, was cor

R:Stoffel, W.; Binczek, E.

Biol. Chem. Hoppe-Seyler 372, 481-488, 1991

A:Title: Transient expression of wild type and mutant human apolipoprotein AI in COS

A:Reference number: S16197; MUID:92029676; PMID:1930731

A:Contents: annotation; extension; studies in reference S02737

R:Stoffel, W.; Krueger, E.; Deutzmann, R.

Hoppe-Seyler's Z. Physiol. Chem. 364, 227-237, 1983

A:Title: Cell-free translation of human liver apolipoprotein AI and AII mRNA processj

A:Reference number: A19913; MUID:83236195; PMID:6407957

A:Accession: B19913

A:Molecule type: protein

A:Residues: 1-6, 'X', '8-13, 'XXX', '17-18, 'XX', '21, 'X', '23-25, 'X', '27-29 <ST2>

R:Enholm, C.; Bozas, S.E.; Tenkano, H.; Kirsbaum, L.; Metso, J.; Murphy, B.; Walke

Biochim. Biophys. Acta 1086, 255-260, 1991

A:Title: The apolipoprotein A-I binding protein of placenta and the SP-40,40 protein

A:Reference number: A56815; MUID:92075698; PMID:1742316

A:Accession: A56815

A:Molecule type: protein

A:Residues: 25-31, 'P', '33 <EHN>

A:Experimental source: serum

A:Note: sequence extracted from NCBI backbone (NCBIRP:69759)

A:Note: 32-Tip was also found

C>Date: 30-Sep-1989 #sequence revision 19-Oct-1995 #text_change 22-Jun-1999

C:Accession: A26529; A26627; S23135; A57766

R:Polites, H.G.; Melchior, G.W.; Castle, C.K.; Marottil, K.R.

Gene 49, 103-110, 1986

A:Title: The primary structure of cynomolgus monkey apolipoprotein A-I deduced from the

A:Reference number: A26529; PMID:87191989; PMID:3106152

A:Accession: A26529

A:Molecule type: mRNA

A:Residues: 1-267 <POL>

A:Cross-references: GB:M15411; NID:q342074; PIDN:AAA36834.1; PID:q342075

R:Herbert, P.N.; Bausserman, L.L.; Lynch, K.M.; Saritelli, A.L.; Kantor, M.A.; Nicolsi,

Biochemistry 26, 1457-1463, 1987

A:Title: Homologues of the human C and A apolipoproteins in the Macaca fascicularis (cyt

A:Reference number: A26627; PMID:87185451; PMID:3105581

A:Accession: A26627

A:Molecule type: protein

A:Residues: 25-48 <HER>

R:Murray, R.W.; Marottil, K.R.

Biochim. Biophys. Acta 1131, 207-210, 1992

A:Title: Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene and corres

A:Reference number: S23135; PMID:92305062; PMID:1610902

A:Accession: S23135

A:Molecule type: DNA

A:Residues: 1-12, 'L', 14-267 <MUR>

A:Cross-references: GB:M83242; NID:q342070; PIDN:AAA36832.1; PID:q342071

R:Sorci-Thomas, M.; Kearns, M.W.

J. Biol. Chem. 266, 18045-18050, 1991

A:Title: Transcriptional regulation of the apolipoprotein A-I gene.

A:Reference number: A57766; PMID:92011532; PMID:1917942

A:Accession: A57766

A:Molecule type: DNA

A:Residues: 1-110 <RES>

A:Cross-references: GB:M69223; NID:q342066; PIDN:AAA36831.1; PID:q553820

C:Comment: The precursor is synthesized in the liver and small intestine. The propeptide

C:Comment: ApoA-I makes up about 70% of the protein of the high density lipoproteins (HDL)

C:Promoting: cholesterol efflux from tissues and by acting as a cofactor for the lecithin

C:Genetics:

A:Insertions: 15/1; 67/2

C:Superfamily: apolipoprotein A-I

C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem repe

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-24/Domain: propeptide #status predicted <PRO>

F:25-267/Product: apolipoprotein A-I #status predicted <MAT>

Query Match 10.0%; Score 187; DB 1; Length 267;

Best Local Similarity 22.5%; Pred. No. 0.00017;

Matches 75; Conservative 61; Mismatches 105; Indels 92; Gaps 11;

5 AAVITMALALSAFSAATQARKGF--DYFSOTSGD-----KGRVQDIN 45

3 ATVTTL--LAVLEP--TGSQARH--FWQODEPQTPMDRVKDLVTYVVALKDSGKDYVSQFE 58

46 QOKKARE--PATLKDSLQDLDNNMKNKFLKRLPLSGSEAPRLPDQVGRMRLOLEEEVK 104

59 GSAIGKQDLNLKLDNNMDSVSTVSKLRQGLPVTQEFMDNLEKTEGRLQSGMSDLEEVK 118

105 ARLOPYMAEAEHELVEGMLRQGLKPYTMDLMEQVALRVOELQRLVVGEDTKAQLLG 164

119 AKVOPYLDDDFQKKWQEEHELVRQKVEPLRKLHETGTRKHLHELKESPLDEEVR---- 173

165 GVDAMALLQGLQSRVYVHTGRFKEFLHPYAESLVSGIGRHOVELHRSVAPHAAPARL 224

174 --DVARA-----HVDALRTHLAPYSDELKQRL 198

225 SRCVQVLSRKLTLLAKALHARIQNDLQRELSRAFAGTGTEGAGDPQMLSEVQR 284

199 AARLEAKENGARLAELHAKASEHLSTLSEKAKPAL-----EDLRQG 241

285 ----LQAFRODTYQLQIAFTPAIDQETEEYQDQ 313

242 LLPVLESFK-----VSFLSALEETVTKLSTQ 267

RESULT 12

apolipoprotein A-I precursor (clone 22ap A1) - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 22-Jun-1999

C:Accession: S06064

R:Parakevoulou, T.B.; Kritis, A.; Zannis, V.

submitted to the EMBL Data Library, July 1989

A:Reference number: S06064

A:Accession: S06064

A:Molecule type: mRNA

A:Residues: 1-266 <PAR>

A:Cross-references: EMBL:X15908; NID:q1457; PIDN:CAA34024.1; PID:q1458

C:Comment: This protein is synthesized in the small intestine.

C:Superfamily: apolipoprotein A-I

C:Keywords: cholesterol metabolism; HDL; intestine; lipid transport; plasma; tandem r

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-24/Domain: propeptide #status predicted <PRO>

F:25-266/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 9.9%; Score 185; DB 1; Length 266;

Best Local Similarity 22.4%; Pred. No. 0.00023;

Matches 61; Conservative 61; Mismatches 122; Indels 28; Gaps 6;

5 AAVITMALALSAFSAATQAR-----KGFMD-----YFSOTSGDKGR--VEQIHQQ 47

3 AAVITLAVL-----TGSQARHFWQODEPRSSMKIKDFATVYVDYVDSGREGVYAOFEAS 59

48 KMAE--PATLKDSLQDLDNNMKNKFLKRLPLSGSEAPRLPDQVGRMRLOLEEEVKAR 106

60 AFGKQDLNLKLDNNMDSVSTVSKLRQGLPVTQEFMDNLEKTEGRLQSGMSDLEEVK 119

107 LQPYMAEAEHELVEGMLRQGLKPYTMDLMEQVALRVOELQRLVVGEDTKAQLLG 166

120 VQPLDFQKKWQEEHELVRQKVEPLRKLHETGTRKHLHELKESPLDEEVR----- 172

167 DEAMALLQGLQSRVYVHTGRFKEFLHPYAESLVSGIGRHOVELHRSVAPHAAPARL 226

173 DSARTHTDTRTKLAPYSDELKQRLAELHETGTRKHLHELKESPLDEEVR 232

227 CVQVLSRKLTLLAKALHARIQNDLQRELS 258

233 ALKDLRQGLLPLVLESFKAQVQNVLDKATKKN 264

apolipoprotein A-I precursor - chicken

C:Species: Gallus gallus (chicken)

C>Date: 31-Mar-1990 #sequence_revision 25-Aug-1995 #text_change 22-Jun-1999

C:Accession: JH0471; A29657; S01453; S28888; A29616; S00187; S10973; S1023; I50156

R:Bhattacharyya, N.; Chattopadhyay, R.; Hirsch, A.; Banerjee, D.

Gene 104, 163-168, 1991

A:Title: Isolation, characterization and sequencing of the chicken apolipoprotein-AI

A:Reference number: JH0471; PMID:92009209; PMID:1916289

A:Accession: JH0471

A:Molecule type: DNA

A:Residues: 1-264 <BHA>

R:Byrnes, L.; Luo, C.C.; Li, W.H.; Yang, C.; Chan, L.

Biochem. Biophys. Res. Commun. 148, 485-492, 1987

A:Title: Chicken apolipoprotein A-I: cDNA sequence, tissue expression and evolution.

A:Reference number: A29657; PMID:88049703; PMID:3118875

A:Accession: A29657

A:Molecule type: mRNA

A:Residues: 1-264 <BYR>

A:Cross-references: EMBL:M17961; NID:q211147; PIDN:AAA48593.1; PID:q211148

R:Rajavashisth, T.B.; Dawson, P.A.; Williams, D.L.; Shackelford, J.E.; Leberitz, H.; I

J. Biol. Chem. 262, 7058-7065, 1987

A:Title: Structure, evolution, and regulation of chicken apolipoprotein A-I.

A:Reference number: S01453; PMID:87223301; PMID:3108248

A:Accession: S01453

A:Molecule type: mRNA

A:Residues: 1-15, '1', '17-147, 'K', '149-264 <RAJ>
A:Cross-references: GB:M25359; EMBL:J02739; NID:g211145; P1DN:AAA48592.1; P1D:g211146
A:Accession: S28888
A:Molecule type: protein
A:Residues: 25-44; 230-256 <RA2>
R:Ferrari, S.; Targui, P.; Drusiani, E.; Calandra, S.; Fregni, M.
Gene 60, 39-46, 1987
A:Title: The complete sequence of chick apolipoprotein A-I mRNA and its expression in the
A:Reference number: A29616; MUID:88152500; PMID:3126099
A:Accession: A29616
A:Molecule type: mRNA
A:Residues: 1-264 <FER>
A:Cross-references: EMBL:M18746; NID:g211149; P1DN:AAA48594.1; P1D:g211150
R:Yang, C.Y.; Gu, Z.W.; Patsch, W.; Weng, S.A.; Kim, T.W.; Chan, L.
FEBS Lett. 224, 261-266, 1987
A:Title: The complete amino acid sequence of proapolipoprotein A-I of chicken high densi
A:Reference number: S00187; MUID:88083548; PMID:3121386
A:Accession: S00187
A:Molecule type: protein
A:Residues: 19-264 <YAN>
R:Banerjee, D.; Mukherjee, T.K.; Redman, C.M.
J. Cell Biol. 101, 1219-1226, 1985
A:Title: Biosynthesis of high density lipoprotein by chicken liver: intracellular transp
A:Reference number: S10973; MUID:86008443; PMID:3930506
A:Accession: S10973
A:Molecule type: protein
A:Residues: 19, 'H', '21-23, 'Q', '25-44 <BAN>
A:Note: the sequence from Fig. 4 is inconsistent with that from Table II in having 41-Th
R:Shackelford, J.E.; Leberer, H.G.
J. Biol. Chem. 258, 7175-7180, 1983
A:Title: Synthesis and secretion of apolipoprotein A-I by chick breast muscle.
A:Reference number: S11023; MUID:83213468; PMID:6406496
A:Accession: S11023
A:Molecule type: protein
A:Residues: 25-44 <SHIA>
R:Lamon-Pava, S.; Ferrari, R.; Rajavashisth, T.B.; Lusis, A.J.; Karathanasis
J. Lipid Res. 33, 831-842, 1992
A:Title: Evolutionary distinct mechanisms regulate apolipoprotein A-I gene expression: D
A:Reference number: I50156; MUID:92381402; PMID:1512510
A:Accession: I50156
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-15, 'I', '17-264 <IAM>
A:Cross-references: GB:M96012; NID:g211158; P1DN:AAA48597.1; P1D:g211159
C:Comment: This protein is synthesized only in the liver in mammals, whereas in chicken
C:Comment: This protein is a major component of the high density lipoproteins in plasma.
C:Genetics:
A:Gene: ApoA1
A:Introns: 15/1; 66/2
C:Superfamily: apolipoprotein A-I
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lip
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-264/Product: apolipoprotein A-I #status experimental <MAT>
Query Match 9.7%; Score 183; DB 1; Length 264;
Best local similarity 21.1%; Pred. No. 0.0003;
Matches 70; Conservative 62; Mismatches 113; Indels 86; Gaps 9;

DB 178 -----HVEELKRNALPYSDLEAOKI. 197
QY 225 SRCVQLSRKLTAKKALHARIQNLDOELRELSAFAGTGTGEEAGDPOMLSEVQR 284
DB 198 SOKLEIREKGIPOQSEYQAKYMEQLSNLREKMT -----PLVGEFRER 240
QY 285 LQAFRODTYIQIAAFTRAIDQETFEVQOOLA 315
DB 241 LTPYENLKNRLISF-----LDELQNSVA 264
RESULT 14
LRRB1B
apolipoprotein A-I precursor (clone pRHA-502) - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text-change 22-Jun-1999
C:Accession: S00230; S20557
R:Pan, T.C.; Hao, Q.L.; Yamin, T.T.; Dai, P.H.; Chen, H.S.; Chen, S.L.; Kroon, P.A.;
Kur, J. Biochem. 170, 99-104, 1987
A:Title: Rabbit apolipoprotein A-I mRNA and gene: evidence that rabbit apolipoprotein
A:Reference number: S00230; MUID:88082866; PMID:3121329
A:Accession: S00230
A:Molecule type: mRNA
A:Residues: 1-265 <PAN>
A:Cross-references: EMBL:X06658; NID:g1461; P1DN:CAA29857.1; P1D:g1462
A:Note: the authors translated the codon AGC for residue 174 as Arg
A:Accession: S20557
A:Molecule type: DNA
A:Residues: 1-17, 'R', '19-44, 'I', '46-122, 'Y', '124-146, 'V', '148-265 <PAN2>
A:Cross-references: EMBL:X06659; NID:g1459; P1DN:CAA29858.1; P1D:g1460
C:Comment: This protein is synthesized in the small intestine.
C:Comment: This protein is a major component of the high density lipoproteins in plas
C:Genetics:
A:Introns: 15/1; 66/2
C:Superfamily: apolipoprotein A-I
C:Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; plas
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-265/Product: apolipoprotein A-I #status experimental <MAT>
Query Match 9.6%; Score 180; DB 1; Length 265;
Best local similarity 21.6%; Pred. No. 0.00046;
Matches 71; Conservative 57; Mismatches 114; Indels 86; Gaps 8;

C>Date: 05-Jan-1996 #sequence-revision 23-Aug-1997 #text-change 13-Aug-1999
 A:Accession: I45853; A56858; A34649
 R:O'Huighin, C.; Chan, L.; Li, M.
 M:1. Biol. Evol. 7, 327-339, 1990
 A:Title: Cloning and sequencing of bovine apolipoprotein A-I cDNA and molecular evolution
 A:Reference number: I45853; M:1. Biol. Evol. 7, 327-339, 1990
 A:Accession: I45853
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-265 <COXH>
 A:Cross-references: GB:M35870; M:1. Biol. Evol. 7, 327-339, 1990
 R:Sparrow, D.A.; Lee, B.R.; Laplaud, P.M.; Aubolton, S.; Baughart, D.; Chapman, M.J.; Gd
 Biochim. Biophys. Acta 1123, 145-150, 1992
 A:Title: Plasma lipid transport in the pre-mature calf, Bos spp: primary structure of h
 A:Reference number: A56858; M:1. Biol. Evol. 7, 327-339, 1990
 A:Accession: A56858
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 19-184, 'QL', 187-265 <SPA>
 A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks
 A:Note: sequence extracted from NCHI backbone (NCBIP-83520)
 R:Aubolton, S.; Sparrow, D.A.; Beaubatie, L.; Baughart, D.; Sparrow, J.T.; Laplaud, P.M.
 Biochem. Biophys. Res. Commun. 166, 833-839, 1990
 A:Title: Characterization and amino-terminal sequence of apolipoprotein AI from plasma h
 A:Reference number: A34649; M:1. Biol. Evol. 7, 327-339, 1990
 A:Accession: A34649
 A:Molecule type: protein
 A:Residues: 25-70 <AUB>
 A:Experimental source: Friesian-Holstein male calves aged 2-4 weeks
 C:Superfamily: apolipoprotein A-I
 C:Keywords: lipid binding; lipoprotein
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-265/Product: apolipoprotein A-I #status experimental <MAT>

Query Match 9.5%; Score 178; DB 2; length 265;
 Best Local Similarity 22.3%; Pred. No. 0.00061;
 Matches 75; Conservative 57; Mismatches 103; Indels 102; Gaps 10;

```

QY 5 AAVLTNALILSAFSAFOAR-----KGFMDYFSOTSQDKGR--VEQIHQ 47
DB 3 AAVLTILAVLFL---TGSQARHPFMOQDDPOSSMDRVKDFATVYVEAIKDSORDVYAQFEAS 59
QY 48 KMARE-PATLKDSLQEDLNMMNKFLKRLRPLSGSEAPRLPDQPYGMRRQLQEELEEVKAR 106
DB 60 ALCKQLNLKLDLMDWDTLASTLSKYVRQLGPTQGFMDNLEKETASLRQEMHKDLEEVKOK 119
QY 107 LQPYMAEAHNLVGMNLEGLRQQLKPYTMDLMPQVALRVQFLQQLRVVCGFTKAQLIGCV 166
DB 120 VQPYLDEFQKKHVEVEIYRQKVAPLGEEFEREGARQKVOELQDKL-----164
QY 167 DEAMALIQCLOSRYVHHTGFKKEI.FHPYAESILVSCIGRHVOELHRSVAPAPASPARLSR 226
DB 165 -----SPLAQELNDRAAHAYETLRQHVAPYS-----190
QY 227 CVOVLSRKLTLKAKAL-----HARIQNLIDLREELSRFACTGTEGAGDPDPOM 276
DB 191 --DDLQRQLRLARLEALKEGGSLEAYHAKASFOELKALGEEK-----AKP-----231
QY 277 LSEEVRORLQAFRODTYLIQIAFTRAIDQETEEVQOO 313
DB 232 VLEDLRQGLPLVLESLKVSILA--AIDEASKKLINAO 265

```

Search completed: January 23, 2003, 17:46:45
 Job time : 39 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2003, 17:34:16 : Search time 20 Seconds

(without alignments)
756.017 Million cell updates/sec

Title: US-09-842-364-3

Perfect score: 1877

Sequence: 1 MASMAAVLTWALALISAFSA.....LWEDITHSLHQGHSLGDP 366

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	350	18.6	382	AP04_PIG	O46409 sus scrofa
2	333	17.7	429	AP04_MACFA	P33621 macaca fasc
3	327	17.4	396	AP04_HUMAN	P06727 homo sapien
4	315	16.8	401	AP04_PAPAN	O28758 papio anubi
5	311.5	16.6	391	AP04_RAT	P02651 rattus norv
6	298.5	15.9	395	AP04_MOUSE	P06728 mus musculu
7	200	10.7	267	AP04_HUMAN	P02647 homo sapien
8	193	10.3	267	AP04_MACFA	P15568 macaca fasc
9	185	9.9	266	AP04_RABIT	P09809 oryctolagus
10	183	9.7	264	AP04_CHICK	P08250 gallus gall
11	179	9.5	262	AP04_ONCMY	O57524 oncorhynch
12	176	9.4	265	AP04_BOVIN	P15497 bos taurus
13	175.5	9.4	266	AP04_CANFA	P02648 canis famli
14	171.5	9.1	264	AP04_CANFA	O42296 anas platyr
15	170	9.1	265	AP04_CANFA	P15497 bos taurus
16	168	9.0	317	AP04_MACFA	P18648 sus scrofa
17	168	9.0	317	AP04_PAPAN	P33298 papio anubi
18	167	8.9	265	AP04_TUPGB	O18759 lumaia gils
19	165	8.8	317	AP04_PIG	P18650 sus scrofa
20	162	8.6	281	AP04_BRARE	O42364 brachydanto
21	161	8.6	264	AP04_MOUSE	O00023 mus musculu
22	160	8.5	262	AP04_SALTR	O91488 salmo trutt
23	156.5	8.3	3674	AP04_HUMAN	O9nc66 homo sapien
24	155	8.3	264	AP04_COTUA	P32918 coturnix co
25	155	8.3	311	AP04_MOUSE	P08826 mus musculu
26	154.5	8.2	312	AP04_RAT	P02650 rattus norv
27	153	8.2	317	AP04_HUMAN	P02649 homo sapien
28	152.5	8.1	298	AP04_CAVPO	P23529 cavia porce
29	152.5	8.1	305	AP04_CANFA	P18649 canis famli
30	152.5	8.1	986	GM13_RAT	O62839 rattus norv
31	150.5	8.0	316	AP04_BOVIN	O03247 bos taurus
32	149	7.9	259	AP04_RAT	P04639 rattus norv
33	149	7.9	311	AP04_RABIT	P18287 oryctolagus

34	148	7.9	262	1	AP01_ONCMY	O57523 oncorhynch
35	143.5	7.6	1940	1	MYH3_HUMAN	P1055 homo sapien
36	141	7.5	3678	1	DMD_MOUSE	P1531 mus musculu
37	139.5	7.4	258	1	AP01_SALSA	P2707 salmo salar
38	138	7.4	1940	1	MYH3_RAT	P12847 rattus norv
39	137	7.3	262	1	AP01_BRARE	O42363 brachydanto
40	136.5	7.3	848	1	MYSP_DIRIM	P13392 drosophila
41	136.5	7.3	1962	1	MYSA_DROME	O97591 drosophila
42	136.5	7.3	3680	1	DMD_CANFA	O97592 canis famli
43	136.5	7.3	3430	1	ACF7_HUMAN	O9upn3 homo sapien
44	136	7.2	1937	1	MYH8_HUMAN	P13535 homo sapien
45	134.5	7.2	3685	1	DMD_HUMAN	P11532 homo sapien

ALIGNMENTS

RESULT 1	ID	AP04_PIG	STANDARD:	PRT:	382 AA.
AC	O46409:	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	Apolipoprotein A-IV precursor (Apo-AIV).				
GN	AP04:				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_Taxid=9823;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Osada J., Iturzaide M., Calleja L., Gonzalez N., Pineiro A.;				
RL	Submitted (DEC-1997) to the EMBL/Genbank/DBD databases.				
CC	-1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND				
CC	CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN				
CC	LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR				
CC	COMPONENT OF HDL AND CHYLOMICRONS.				
CC	-1- SUBCELLULAR LOCATION: Extracellular.				
CC	-1- TISSUE SPECIFICITY: SECRETED IN PLASMA.				
CC	-1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH				
CC	22-MR IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED) 11-				
CC	MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-				
CC	HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY				
CC	THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL				
CC	ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.				
CC	-1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: AJ222966; CA11020.1; -				
DR	HSSP: P32851; IBR0.				
DR	Interpro: IPR000074; Apolipoprotein.				
DR	Plasma; lipid transport; HDL; Chylomicron; Repeat; Signal.				
KW	Plasma; lipid transport; HDL; Chylomicron; Repeat; Signal.				
FT	SIGNAL	1	20		BY SIMILARITY.
FT	CHAIN	21	382		APOLIPOPROTEIN A-IV.
FT	DOMAIN	31	330		13 X 22 AA APPROXIMATE TANDEM REPEATS.
FT	REPEAT	33	34		1.
FT	REPEAT	60	81		2.
FT	REPEAT	82	103		3.
FT	REPEAT	115	136		4.
FT	REPEAT	137	158		5.
FT	REPEAT	159	180		6.
FT	REPEAT	181	202		7.
FT	REPEAT	203	224		8.
FT	REPEAT	225	246		9.

CC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC TISSUE=Intestine; PubMed-2930771;
 RX MEDLINE-89194198;
 RA Yang C., Gu Z.M., Xiong W., Rosseneu M., Yang H.X., Lee B.M.,
 RA Goto A.M., Jr., Chan L.,
 RT "The primary structure of human apolipoprotein A-IV.";
 RL Biochim. Biophys. Acta 1002:231-237(1989).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87041474; PubMed-3095836;
 RA Karathanasis S.K., Oeltgen P., Haddad I.A., Antonarakis S.E.;
 RT "Structure, evolution, and polymorphisms of the human apolipoprotein
 A4 gene (APOA4).";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8457-8461(1986).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86296629; PubMed-3755616;
 RA Karathanasis S.K., Yunis I.;
 RT "Structure, evolution, and tissue-specific synthesis of human
 apolipoprotein AIV.";
 RL Biochemistry 25:3962-3970(1986).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87250378; PubMed-3036793;
 RA Elshourbagy N.A., Walker D.W., Patk Y.K., Hoguski M.S., Freeman M.,
 RA Gordon J.I., Taylor J.M.;
 RT "Structure and expression of the human apolipoprotein A-IV gene.";
 RL J. Biol. Chem. 262:7973-7981(1987).
 RN [15]
 RP SEQUENCE OF 21-396 FROM N.A.
 RX MEDLINE-86111885; PubMed-3080432;
 RA Elshourbagy N.A., Walker D.W., Hoguski M.S., Gordon J.I., Taylor J.M.;
 RT "The nucleotide and deduced amino acid sequence of human
 apolipoprotein A-IV mRNA and the close linkage of its gene to the
 genes of apolipoproteins A-I and C-III.";
 RL J. Biol. Chem. 261:1998-2002(1986).
 RN [16]
 RP SIGNAL SEQUENCE CLEAVAGE SITE.
 RX MEDLINE-84161950; PubMed-6706947;
 RA Gordon J.I., Bisgaier C.L., Sims H.F., Sachdev O.P., Glickman R.M.,
 RA Strauss A.W.;
 RT "Biosynthesis of human preapolipoprotein A-IV.";
 RL J. Biol. Chem. 259:468-474(1984).
 RN [17]
 RP REVIEW ON POLYMORPHISM.
 RA Lohse P., Brewer H.B. Jr.;
 RT "Genetic polymorphism of apolipoprotein A-IV.";
 RL Curr. Opin. Lipidol. 2:90-95(1991).
 RN [18]
 RP VARIANT A-IV*2.
 RX MEDLINE-90277616; PubMed-2351649;
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;
 RT "Genetic polymorphism of human plasma apolipoprotein A-IV is due to
 nucleotide substitutions in the apolipoprotein A-IV gene.";
 RL J. Biol. Chem. 265:10061-10064(1990).
 RN [19]
 RP VARIANTS A-IV*0 AND A-IV*3.
 RX MEDLINE-90324273; PubMed-1973689;
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;
 RT "Human plasma apolipoproteins A-IV-0 and A-IV-3. Molecular basis for
 two rare variants of apolipoprotein A-IV-1.";
 RL J. Biol. Chem. 265:12734-12739(1990).
 RN [10]
 RP VARIANTS.
 RX MEDLINE-91310615; PubMed-1677358;
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;
 RT "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-
 1(Thr-347->Ser), apoA-IV-0(lys-167->Gln, Gln-360->His), and apoA-IV-
 3(Gln-165->Lys).";
 RL J. Biol. Chem. 266:13513-13518(1991).
 RN [11]
 RP ERRATUM.
 RA Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;
 RL J. Biol. Chem. 266:19866-19866(1991).
 RN [12]
 RP VARIANT MET-13.
 RX MEDLINE-92238494; PubMed-1349197;
 RA von Eckardstein A., Funke H., Schulte M., Erben M., Schulte H.,
 RA Assmann G.;
 RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV
 gene are associated with changes in the concentration of apo B- and
 apo A-I-containing lipoproteins in a normal population.";
 RL Am. J. Hum. Genet. 50:1115-1128(1992).
 RN [13]
 RP VARIANT SER-147.
 RX MEDLINE-92144647; PubMed-1737067;
 RA Tenkanen H., Koskinen P., Melsö J., Haumann M., Jukka M.,
 RA Kauppinen-Makelin R., Kontula K., Taskiran M.R., Manttari M.,
 RA Manninen V., Ehnholm G.;
 RT "A novel polymorphism of apolipoprotein A-IV is the result of an
 asparagine to serine substitution at residue 127.";
 RL Biochim. Biophys. Acta 1138:27-33(1992).
 RN [14]
 RP VARIANT A-IV*5.
 RX MEDLINE-93138374; PubMed-1487136;
 RA Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,
 RA Perrelli R.E., Pollitzer W.S.;
 RT "Molecular basis of a unique African variant (A-IV 5) of human
 apolipoprotein A-IV and its significance in lipid metabolism.";
 RL Genet. Epidemiol. 9:379-388(1992).
 RN [15]
 RP VARIANT BUDAPEST-2 LYS-44; BUDAPEST-1 CYS-305 AND SER-367.
 RX MEDLINE-95245341; PubMed-7728150;
 RA Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,
 RA Caszlar A.;
 RT "Apolipoprotein A-IV polymorphism in the Hungarian population: gene
 frequencies, effect on lipid levels, and sequence of two new
 variants.";
 RL Hum. Mutat. 5:58-65(1995).
 RN [16]
 RP VARIANTS FCHE, SHAPLE SER-161; LFU-178 AND GLN-264.
 RX MEDLINE-97114287; PubMed-8956036;
 RA Deeb S.S., Nevlin D.N., Iwasaki L., Brunzell J.D.;
 RT "Two novel apolipoprotein A-IV variants in individuals with familial
 combined hyperlipidemia and diminished levels of lipoprotein lipase
 activity.";
 RL Hum. Mutat. 8:319-325(1996).
 CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
 CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
 LIPASE BY APOC-II: POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
 COMPONENT OF HDL AND CHYLOMICRONS.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.
 CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED II-
 MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
 HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
 THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECITHIN:CHOLESTEROL
 ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
 CC -1- POLYMORPHISM: EIGHT ALLELES HAVE BEEN CHARACTERIZED (APOA-IV-0
 TO APOA-IV*7). APOA-IV*1 IS THE MAJOR ALLELE (90%), IV*2 IS ALSO
 COMMON (8%), THE OTHERS ARE RARE ALLELES.
 CC -1- DISEASE: DEFECTS IN APOA4 MAY CONTRIBUTE, ALONG WITH DEFECTS IN
 OTHER GENES OR ENVIRONMENTAL FACTORS, TO THE DEVELOPMENT OF
 FAMILIAL COMBINED HYPERLIPIDEMIA (FCHL).
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

OY 280 EVRORLQAFRODTYLOIAAFTRALIDETEVQOOLAPPPGHSAPAFEPQOTDSGV--- 336
 DB 295 QLDQOYEVERRAVERPLGDFNFALVQOMEKFRQOLG-----SSGDVESH 339
 OY 337 LSKLQARLDLWEDITHSLHDG 359
 DB 340 LSFLEKRLREKVSFMSFTLQKKG 362

RESULT 6
 APOA4_MOUSE STANDARD; PRT; 395 AA.
 AC P06728;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Apolipoprotein A-IV precursor (Apo-AIV).
 GN APOA4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87089722; PubMed=3796595;
 RA Williams S.C., Bruckheimer S.M., Lusis A.J., Lehoucq R.C.,
 RA Kinniburgh A.J.;
 RT "Mouse apolipoprotein A-IV gene: nucleotide sequence and induction by
 a high-lipid diet."
 RL Mol. Cell. Biol. 6:3807-3814(1986).
 RN (2)
 RP REVISIONS.
 RA Kinniburgh A.J.;
 RL Submitted (DEC-1986) to the EMBL/Genbank/DBD databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J, and 129/J;
 RX MEDLINE=91286309; PubMed=1648102;
 RA Reue K., Leete T.H.;
 RT "Genetic variation in mouse apolipoprotein A-IV due to insertion and
 deletion in a region of tandem repeats."
 RL J. Biol. Chem. 266:12715-12721(1991).
 CC -I- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
 CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
 LIPASE BY APOC-II. POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
 COMPONENT OF HDL AND CHYLOMICRONS.
 CC -I- SUBCELLULAR LOCATION: Extracellular.
 CC -I- TISSUE SPECIFICITY: SECRETED IN PLASMA.
 CC -I- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDDEM REPEATS (EACH
 22-MER IS ACTUALLY A TANDDEM ARRAY OF TWO, A AND B, RELATED 11-
 MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
 HELICAL, AND MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
 THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECTIHIN:CHOLESTEROL
 ACUTRANSFERASE (LCAT) ACTIVATING ABILITIES.
 CC -I- POLYMORPHISM: THERE IS A POLYMORPHISM WITHIN A SERIES OF IMPERFECT
 REPEATS ENCODING THE SEQUENCE E-Q-[AV]-Q. INSERTIONS OR DELETIONS
 OF 12 NUCLEOTIDES HAVE GIVEN RISE TO THREE FORMS CHARACTERIZED BY
 THREE (129), FOUR (C57BL/6), OR FIVE (M.CASTANEUS) COPIES OF THE
 REPEAT UNIT.
 CC -I- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: M13966; AAA37253.1;
 DR EMBL: M64249; AAA37215.1;
 DR EMBL: M64248; AAA37214.1;
 DR PIR: A25281; A25281.

DR MGD; MGI:88051; APOA4.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein; 2.
 KW Plasma; Lipid transport; HDL; Chylomicron; Repeat; Signal;
 KM Polymorphism.

FT	CHAIN	1	20	APOLIPOPROTEIN A-IV
FT <td>SIGNAL</td> <td>21</td> <td>395</td> <td>13 X 22 AA APPROXIMATE TANDDEM REPEATS.</td>	SIGNAL	21	395	13 X 22 AA APPROXIMATE TANDDEM REPEATS.
FT <td>DOMAIN</td> <td>33</td> <td>330</td> <td></td>	DOMAIN	33	330	
FT <td>REPEAT</td> <td>33</td> <td>54</td> <td>1.</td>	REPEAT	33	54	1.
FT <td>REPEAT</td> <td>60</td> <td>81</td> <td>2.</td>	REPEAT	60	81	2.
FT <td>REPEAT</td> <td>82</td> <td>103</td> <td>3.</td>	REPEAT	82	103	3.
FT <td>REPEAT</td> <td>115</td> <td>136</td> <td>4.</td>	REPEAT	115	136	4.
FT <td>REPEAT</td> <td>137</td> <td>158</td> <td>5.</td>	REPEAT	137	158	5.
FT <td>REPEAT</td> <td>159</td> <td>180</td> <td>6.</td>	REPEAT	159	180	6.
FT <td>REPEAT</td> <td>181</td> <td>202</td> <td>7.</td>	REPEAT	181	202	7.
FT <td>REPEAT</td> <td>203</td> <td>224</td> <td>8.</td>	REPEAT	203	224	8.
FT <td>REPEAT</td> <td>225</td> <td>246</td> <td>9.</td>	REPEAT	225	246	9.
FT <td>REPEAT</td> <td>247</td> <td>268</td> <td>10.</td>	REPEAT	247	268	10.
FT <td>REPEAT</td> <td>269</td> <td>286</td> <td>11.</td>	REPEAT	269	286	11.
FT <td>REPEAT</td> <td>287</td> <td>308</td> <td>12.</td>	REPEAT	287	308	12.
FT <td>REPEAT</td> <td>309</td> <td>330</td> <td>13.</td>	REPEAT	309	330	13.
FT <td>DOMAIN</td> <td>366</td> <td>389</td> <td>GLU/GLN-RICH.</td>	DOMAIN	366	389	GLU/GLN-RICH.
FT <td>VARIANT</td> <td>382</td> <td>385</td> <td>MISSING (IN SOME STRAINS).</td>	VARIANT	382	385	MISSING (IN SOME STRAINS).
FT <td>CONFLICT</td> <td>15</td> <td>15</td> <td>0 -> K (IN REF. 1).</td>	CONFLICT	15	15	0 -> K (IN REF. 1).
FT <td>CONFLICT</td> <td>63</td> <td>63</td> <td>E -> R (IN REF. 1).</td>	CONFLICT	63	63	E -> R (IN REF. 1).
FT <td>CONFLICT</td> <td>207</td> <td>207</td> <td>S -> A (IN REF. 1).</td>	CONFLICT	207	207	S -> A (IN REF. 1).
FT <td>CONFLICT</td> <td>288</td> <td>288</td> <td>RQ -> KA (IN REF. 1).</td>	CONFLICT	288	288	RQ -> KA (IN REF. 1).
FT <td>CONFLICT</td> <td>294</td> <td>295</td> <td>NK -> GG (IN REF. 1).</td>	CONFLICT	294	295	NK -> GG (IN REF. 1).
FT <td>CONFLICT</td> <td>315</td> <td>316</td> <td></td>	CONFLICT	315	316	
SO <td>SEQUENCE</td> <td>395 AA;</td> <td>45029 MW;</td> <td>5FE27D023626257 CRC64;</td>	SEQUENCE	395 AA;	45029 MW;	5FE27D023626257 CRC64;

Query Match 15.9%; Score 298.5; DB 1; Length 395;
 Best Local Similarity 24.4%; Pred. No. 2,5e-11;
 Matches 91; Conservative 85; Mismatches 164; Indels 33; Gaps 7;

OY 5 AAVLTALALISA-----FSATQARKGFMDYPSQTSQD-KGRVQDHOOKAREPATL-KD 58
 DB 5 AAVLTALALVAITGTRAVTSDQVANYVMDYFTQLSNNAKAEVPOFTDVTQOLSTLFOD 64
 OY 59 SLEODLNNKKELEKRLPLSGSEAPRLPDPVGRQLOLEELVKARLOPYAAEAHELV 118
 DB 65 KLDIASTYADGVNKKLVPPFVQVLSGLHAKETPHVKKFKEKLELDLNDMMHPANKVTQTF 124
 OY 119 GNNLEGLRQOLKRYTMDIMQVALLRQVLEOQLKVGQEDTKAOLGCVD---EANAALQ 174
 DB 125 GNNQKLOEHLKPYAVDLOQINTQOEMKLOLTPYIQRTTIKEVNDLHNSMPLAT 184
 OY 175 GLOSRVYVNHTRGPKKLEPHYPAESIVSGICRHVOHLSHVAHPAPASPARISRCVOYLSRK 234
 DB 185 NIKDKRNKNMNEHLKCHLTTPRANLTKATIIQNLIDLRKSLAPLTVGVQEKINHOMEGIAFO 244
 OY 235 LTKAKALARIQONLDQLEELSRFACT-----GTEGAGDPQWLSEEVRORLQAFR 289
 DB 245 MKRNAEELQTKVSAKIDOLQKNLAPLVEDQSVKQNTGELQSLSDLNQLOQYEFRR 304
 OY 290 QUDYVLOIAAFTRALIDETEVQOOLAPPPGHSAPAFEPQOTDSGV---LSKLQARLD 346
 DB 305 RYVEPGEEMNKALVQOLEFROOLGP-----NSGEVSHLSLEKSLRE 349
 OY 347 LWEDITHSLHDG 359
 DB 350 KVNVSFMSLEKKG 362

RESULT 7
 APOA1_HUMAN STANDARD; PRT; 267 AA.
 AC P02647;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 GN APOA1.

- OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Furelestoml;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI.TaxID=9606.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84221405; PubMed=6328445;
 RA Sharpe C.R., Sidoli A., Shelley C.S., Lucero M.A., Shoulders C.C.,
 Baralle F.E.;
 RT "Human apolipoproteins AI, AII, CII and CIII. cDNA sequences and mRNA
 abundance.";
 RL Nucleic Acids Res. 12:3917-3932(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85026665; PubMed=6207999;
 RA Sellmeyer J.J., Protter A.A., Frossard P., Levy-Wilson B.;
 RT "Isolation and DNA sequence of full-length cDNA and of the entire
 gene for human apolipoprotein AI -- discovery of a new genetic
 polymorphism in the apo AI gene.";
 RL DNA 3:309-317(1984).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83220822; PubMed=6406984;
 RA Shoulders C.C., Kornblith A.R., Munro B.S., Baralle F.E.;
 RT "Gene structure of human apolipoprotein AI.";
 RL Nucleic Acids Res. 11:2827-2837(1983).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83220772; PubMed=6304641;
 RA Cheung P., Chan L.;
 RT "Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.";
 RL Nucleic Acids Res. 11:3703-3715(1983).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84119464; PubMed=6198645;
 RA Law S.W., Brewer H.B. Jr.;
 RT "Nucleotide sequence and the encoded amino acids of human
 apolipoprotein A-I mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:66-70(1984).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86008382; PubMed=2995392;
 RA Law S.W., Brewer H.B. Jr.;
 RT "Tangier disease. The complete mRNA sequence encoding for
 preproapo-A-I.";
 RL J. Biol. Chem. 260:12810-12814(1985).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84016011; PubMed=6413973;
 RA Karathanasis S.K., Zannis V.I., Breslow J.L.;
 RT "Isolation and characterization of the human apolipoprotein A-I
 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6147-6151(1983).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89377481; PubMed=2673706;
 RA Mogulievsky N., Roobol C., Loriau R., Guillaume J.P., Jacobs P.,
 Cravador A., Herzog A., Brouwers L., Scarso A., Gilles P.,
 Holmquist L., Carlson L.A., Bollen A.;
 RT "Production of human recombinant proapolipoprotein A-I in Escherichia
 coli: purification and biochemical characterization.";
 RL DNA 8:429-436(1989).
 RN [9]
 RP SEQUENCE FROM N.A. (VARIANT TANGIER).
 RX MEDLINE=88196137; PubMed=3129297;
 RA Medrides S.C., Ruiz-Opazo N., Hayden M., Nussbaum A.L., Breslow J.L.,
 Zannis V.I.;
 RT "Sequence and expression of Tangier apoA-I gene.";
 RL Eur. J. Biochem. 173:465-471(1988).
 RN [10]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Strausberg R.;
 RN Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 RN [11]
 RP SEQUENCE OF 118-267 FROM N.A.
 RX MEDLINE=83091059; PubMed=6294659;
 RA Breslow J.L., Ross D., McPherson J., Williams H.W., Kurnit D.,
 Nussbaum A.L., Karathanasis S.K., Zannis V.I.;
 RT "Isolation and characterization of cDNA clones for human
 apolipoprotein A-I.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6861-6865(1982).
 RN [12]
 RP SEQUENCE OF 19-27.
 RX MEDLINE=83256553; PubMed=6409108;
 RA Brewer H.B. Jr., Fairwell T., Kay L., Meng M., Ronan K., Law S.,
 Light J.A.;
 RT "Human plasma proapoA-I: isolation and amino-terminal sequence.";
 RL Biochem. Biophys. Res. Commun. 113:626-632(1983).
 RN [13]
 RP SEQUENCE OF 25-267.
 RX MEDLINE=78123731; PubMed=204308;
 RA Brewer H.B. Jr., Fairwell T., Larue A., Ronan K., Houser A.,
 Bronzert T.J.;
 RT "The amino acid sequence of human APOA-I, an apolipoprotein isolated
 from high density lipoproteins.";
 RL Biochem. Biophys. Res. Commun. 80:623-630(1978).
 RN [14]
 RP SEQUENCE OF 25-267.
 RX MEDLINE=75133493; PubMed=164450;
 RA Baker H.N., Gotto A.M. Jr., Jackson R.L.;
 RT "The primary structure of human plasma high density apolipoprotein
 glutamine I (APOA-I). II. The amino acid sequence and alignment of
 cyanogen bromide fragments IV, III, and I.";
 RL J. Biol. Chem. 250:2725-2738(1975).
 RN [15]
 RP SEQUENCE OF 25-56.
 RX MEDLINE=88331387; PubMed=3047170;
 RA Yui Y., Aoyama T., Morishita H., Takahashi M., Takatsu Y., Kawai C.;
 RT "Serum prostacyclin stabilizing factor is identical to apolipoprotein
 A-I (APO A-I). A novel function of Apo A-I.";
 RL J. Clin. Invest. 82:803-807(1988).
 RN [16]
 RP SEQUENCE OF 25-48.
 RX MEDLINE=89380318; PubMed=2506184;
 RA Manjunath P., Marcel Y.L., Uma J., Seidah N.G., Chretien M.,
 Chapdelaine A.;
 RT "Apolipoprotein A-I binds to a family of bovine seminal plasma
 proteins.";
 RL J. Biol. Chem. 264:16853-16857(1989).
 RN [17]
 RP SEQUENCE OF 25-43.
 RX MEDLINE=88070603; PubMed=3120314;
 RA Pifoll R.P., Ordovas J.M., Rosenberg I., Schaeffer E.J.,
 Pereira M.E.A.;
 RT "Similarity of cruzin, an inhibitor of Trypanosoma cruzi
 neuraminidase, to high-density lipoprotein.";
 RL Science 238:1417-1419(1987).
 RN [18]
 RP SEQUENCE OF 25-42.
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 1994.";
 RL Electrophoresis 15:1459-1465(1994).
 RN [19]
 RP PALMITOYLATION.
 RX MEDLINE=86140194; PubMed=3005308;
 RA Hoeg J.M., Meng M.S., Ronan R., Fairwell T., Brewer H.B. Jr.;
 RT "Human apolipoprotein A-I. Post-translational modification by fatty
 acid acylation.";
 RL J. Biol. Chem. 261:3911-3914(1986).
 RN [20]
 RP PROCESSING.
 RX MEDLINE=83195100; PubMed=6405383;

RA Zannis V.I., Karathanasis S.K., Keutmann H.T., Goldberger G.,
 Breslow J.L., "Intracellular and extracellular processing of human apolipoprotein
 A-I: secreted apolipoprotein A-I isoprotein 2 is a propeptide.",
 Proc. Natl. Acad. Sci. U.S.A. 80:2574-2578(1983).
 (21)
 RN STRUCTURE BY NMR OF 190-209.
 RX MEDLINE-96270776; PubMed-8664326;
 RA Wang G., Treleaven W.D., Cushley R.J.;
 KT "Conformation of human serum apolipoprotein A-I (166-185) in the
 RT presence of sodium dodecyl sulfate or dodecylphosphocholine by 1H-NMR
 and CD. Evidence for specific peptide-SDS interactions.",
 RL Biochim. Biophys. Acta 1301:174-184(1996).
 (22)
 RN X-RAY CRYSTALLOGRAPHY (4.0 ANGSTROMS) OF 67-267.
 RX MEDLINE-98024124; PubMed-9356442;
 RA Borthan D.W., Rogers D.P., Knigler J.A., Brouillette C.G.;
 RT "Crystal structure of truncated human apolipoprotein A-I suggests a
 RL lipid-bound conformation.",
 Proc. Natl. Acad. Sci. U.S.A. 94:12291-12296(1997).
 (23)
 RN VARIANT MILANO.
 RX MEDLINE-83109095; PubMed-6401735;
 RA Weisgraber K.H., Rall S.C. Jr., Bersot T.P., Mahley R.W.,
 RT "Apolipoprotein A-I Milano. Detection of normal A-I in affected
 RT subjects and evidence for a cysteine for arginine substitution in the
 RT variant A-I.",
 RL J. Biol. Chem. 258:2508-2513(1983).
 (24)
 RN VARIANT TANGIER.
 RX MEDLINE-83300108; PubMed-6412234;
 RA Schmitz G., Assmann G., Rall S.C. Jr., Mahley R.W.;
 RT "Tangier disease: defective recombination of a specific Tangier
 RT apolipoprotein A-I isoform (pro-apo A-I) with high density
 RT lipoproteins.",
 RL Proc. Natl. Acad. Sci. U.S.A. 80:6081-6085(1983).
 (25)
 RN VARIANT NORMAN.
 RX MEDLINE-84289383; PubMed-6432779;
 RA Rall S.C. Jr., Weisgraber K.H., Mahley R.W., Ogawa Y., Fielding C.J.,
 RT "Abnormal lecithin:cholesterol acyltransferase activation by a human
 RT apolipoprotein A-I variant in which a single lysine residue is
 RT deleted.",
 RL J. Biol. Chem. 259:10063-10070(1984).
 (26)
 RN

Query Match 10.7%; Score 200; DB 1; Length 267;
 Best Local Simlarity 22.2%; Pred. No. 1.3e-05;
 Matches 74; conservative 61; Mismatches 106; Indels 92; Gaps 10;

QY 5 AAVLTMAALISAFSATQAKKFWD-----YFSQTSCKGR--VEQIH 45
 DB 3 AAVLTMAVLEL--TGSQARH-FWQDEPQSWDRVKDLATVYDVYLKSGRDYVSQFE 58
 QY 46 OQKMARE-PATLKDSLIEDIINNKKFLKRLPLSGSEAPRLPODPYQMRLOEELBEVK 104
 DB 59 GSAIGQOLNKLKLDNDVSTSTSKLRQIGPTQOEFWNLKRYTEGLKQEMSKDLEEVK 118
 QY 105 ARLQPYMAENHELVGNLBSGLRQOLKPYTMDLEQVALRYVELQEDQLRVVGEDTKAQLG 164
 DB 119 AKQAPYLDQFOKKQWQEMELRYOKVRLRAELOEGAROKLHELOEKLSPIGSEMR---- 173
 QY 165 GVEAVALLQGLQSVVNHNGRKHLFHYAESELVSGICGHVQELIRSVAPARAPARL 224
 DB 174 --DRAA-----HVDALHTLAPYSDELRLQRL 198
 QY 225 SRGVYVLRKTLKAKALHARLQONLDQREELSRFAFGTGEAGAPRQMLSEEVQR 284
 DB 199 AARLEALKENGARLAEYNAKATEHLSTLSEKKKRAL-----EDLRQG 241
 QY 285 ----LDAFRODYLYQIAATFRAIDQTEVEYQOO 313

DB 242 LLPVLESFK-----VSFLSALEETKKNFTQ 267
 ID APAL_MACFA STANDARD; PRT: 267 AA.
 AC P15568; P17929;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 GN APOA1.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
 OS Papio hamadryas (Hamadryas baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID:9541, 9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 KC SPECIES-M.fascicularis;
 RX MEDLINE-87191989; PubMed-3106152;
 RA Polites H.G., Melchior G.W., Castile C.R., Marolli K.R.;
 RT "The primary structure of cynomolgus monkey apolipoprotein A-I
 RT deduced from the cDNA sequence: comparison to the human sequence.",
 RL Gene 49:103-110(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 KC SPECIES-M.fascicularis;
 RX MEDLINE-92305062; PubMed-1610902;
 RA Murray R.W., Marolli K.R.;
 RT "Nucleotide sequence of the cynomolgus monkey apolipoprotein A-I gene
 RT and corresponding flanking regions.",
 RL Biochim. Biophys. Acta 1131:207-210(1992).
 RN [3]
 RP SEQUENCE OF 25-48.
 KC SPECIES-M.fascicularis;
 RX MEDLINE-87185451; PubMed-3105581;
 RA Herbert P.N., Hausserman J.L., Lynch K.M., Sartelli A.L.,
 RT Kantor M.A., Nicotoli R.J., Shulman R.S.;
 RT "Homologues of the human C and A apolipoproteins in the Macaca
 RT fascicularis (cynomolgus) monkey.",
 RL Biochemistry 26:1457-1463(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 KC SPECIES-P.hamadryas; TISSUE-Liver;
 RX MEDLINE-89232739; PubMed-2907746;
 RA Hixson J.E., Borenstein S., Cox L.A., Rainwater D.L., Vandeberg J.L.;
 RT "The baboon gene for apolipoprotein A-I: characterization of a cDNA
 RT clone and identification of DNA polymorphisms for genetic studies of
 RT cholesterol metabolism.",
 RL Gene 74:483-490(1988).
 RN [5]
 RP SEQUENCE OF 1-10 FROM N.A.
 RC SPECIES-M.fascicularis;
 RA Sorci-Thomas M.;
 RT Submitted (Oct-1991) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COPACITOR FOR
 CC THE LECTININ CHOLESTEROL ACYLTRANSFERASE (LCAT).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
 CC CHROMIDROMS.
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@db-sib.ch.

DR EMBL: M15411; AAA36834.1; -
 DR EMBL: M83242; AAA36832.1; -
 DR EMBL: M35634; AAA35380.1; -
 DR EMBL: M69223; AAA36831.1; -
 DR PIR: A26529; A26529.
 DR PIR: A26627; A26627.
 DR PIR: JS0079; JS0079.
 DR PIR: S23135; S23135.
 DR HSSP: P02647; IGW4.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein; 1.
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.

FT STGNL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 267
 FT DOMAIN 68 267
 FT REPEAT 68 89
 FT REPEAT 90 111
 FT REPEAT 112 122
 FT REPEAT 123 144
 FT REPEAT 145 166
 FT REPEAT 167 188
 FT REPEAT 189 210
 FT REPEAT 211 232
 FT REPEAT 233 243
 FT REPEAT 244 267
 FT CONFLICT 13 13
 SQ SQUINCE 267 AA; 30735 MW; 869955C0408E21 CRC64;

Query Match 10.3%; Score 193; DB 1; Length 267;
 Best Local Similarity 22.2%; Pred. No. 3.5e-05;
 Matches 74; Conservative 61; Mismatches 106; Indels 92; Gaps 10;

QY 5 AAVLTALALISAFSAOARKFW--DYFSOTSD-----KGRVEQIN 45
 DB 3 ATVTLLVLFV--TGSQARH-FWQODEPQTPMDRVKDLVTYVEALKDQSKDYVSQFE 58
 QY 46 QOKKARE-PATIKDSDODNNMKNFKLRLPSGSEAPRLPPDPCVGRROLODELEPVK 104
 DB 59 GSAIGKQDNLKLDNMWDSVSTVSKLRKQDLPVTOEFTWMDNLEKEEGIKQMSKDLSEVK 118
 QY 105 ARLQPYAAEAHELWGMVFGILRQQLPYTMDLMFOVALRVOELOEOI.RVNGEDTKAQLIG 164
 DB 119 AKVQPYDDPQKKQEMELRYRQKVEPLRALHELHGTROKLEHEKSLPSGEEVR----- 173
 QY 165 GVDEAMAILQCLQSRVYVHTGRFKEIHPYAFSLVSGIGRHVOELHNSVAIHAPASPAKI 224
 DB 174 --DRAA-----HVDALRTYHIALYSDLEIKQRI 198
 QY 225 SRCQVLSRKLTAKKALLIRIOONLDOLRELSRAFAVGTEGACGPDQMLSEEVOR 284
 DB 199 AARLEALKENCGARLALEYHAASSEHISTLSEKAPAL-----EDLNQG 241
 QY 285 ---LQAFKQDTYLIQIAFTRAIDQETFEYQQ 313
 DB 242 LLPVLESFK-----VSFLSALEETKLTSTQ 267

RESULT 9
 APAL_RABIT STANDARD; PRT; 266 AA.
 ID APAL_RABIT
 AC P09809;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 GN APOA1.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NC MBL:TaxID=9886;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-22AP AI; TISSUE=Small intestine;
 RA Paraskevopoulos T.B., Krilis A., Zannis V.I.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBD databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Intestine;
 RA MEDLINE=88082866; PubMed=3121329;
 RA Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,
 RA Kroon P.A., Chao Y.S.;
 RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit
 RT apolipoprotein A-I is synthesized in the intestine but not in the
 RT liver";
 RL Eur. J. Biochem. 170:99-104(1987).
 RN [3]
 RP SEQUENCE OF 25-266.
 RX MEDLINE=87030294; PubMed=3095115;
 RA Yang C., Yang T., Pomall H.J., Gotto A.M. Jr.;
 RT "The primary structure of apolipoprotein A-I from rabbit high-density
 RT lipoprotein";
 RL Eur. J. Biochem. 160:427-431(1986).
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
 CC CHYLOMICRONS.
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@db-sib.ch).

CC EMBL: X15908; CAA34024.1; -
 DR EMBL: X06558; CAA29857.1; -
 DR EMBL: X06559; CAA29858.1; -
 DR PIR: S06064; LPRB12.
 DR PIR: S00230; LPRB1B.
 DR PIR: A24998; A24998.
 DR HSSP: P02647; IAV1.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein; 1.
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.

FT PROPEP 19 24
 FT CHAIN 25 266
 FT DOMAIN 67 266
 FT REPEAT 67 88
 FT REPEAT 89 110
 FT REPEAT 111 121
 FT REPEAT 122 143
 FT REPEAT 144 165
 FT REPEAT 166 187
 FT REPEAT 188 209
 FT REPEAT 210 231
 FT REPEAT 232 242
 FT REPEAT 243 266
 FT CONFLICT 18 18
 FT CONFLICT 44 44
 FT CONFLICT 45 45
 FT CONFLICT 107 107
 FT CONFLICT 123 123
 FT CONFLICT 147 147
 FT CONFLICT 150 150
 FT CONFLICT 191 191
 FT CONFLICT 195 195
 FT CONFLICT 211 211

A -> R (IN REF. 2; CAA29858).
 MISSING (IN REF. 3).
 V -> I (IN REF. 2; CAA29858).
 E -> Q (IN REF. 3).
 Y -> F (IN REF. 2; CAA29857).
 A -> V (IN REF. 2; CAA29858 AND 3).
 R -> G (IN REF. 2; CAA29858/CAA29857).
 N -> Q (IN REF. 3).
 MISSING (IN REF. 2; CAA29858/CAA29857).
 S -> K (IN REF. 2; CAA29858/CAA29857).

FT CONFLICT 255 256 VL -> LV (IN REF. 2)
 FT CONFLICT 256 256 L -> V (IN REF. 2; CAA29858/CAA29857).
 SO SOURCE 256 AA: 30591 MW: 0FF6D386497C7D2 CRC64:
 Query Match 9.9%; Score 185; DB 1; Length 266;
 Best Local Similarity 22.4%; Pred. No. 0.0001;
 Matches 61; Conservative 61; Mismatches 122; Indels 28; Gaps 6;

QY 5 AAVLTWALALISAFSAIOAR-----KCFWD-----YFSQSGDKCR--VQIHOQ 47
 DB 3 AVVLTWALVFL--TGSOARHFWQDEPRSSWDKIKDPATVYVDIVKSGREYVAQFEAS 59
 QY 48 KWARE-PATIKDLEQDINNMMKFELEKRLPLSGSEAPRLPODPVGMRRLOLELEEVKAR 106
 DB 60 AFGKOLNKLIDNMOSLSSTVSKLOEQLOIPYOFHWDLNLEKTEGILKEPMKNDIQEVROK 119
 QY 107 IQPYAAEAHELVGNMIFGLKQOLKPYTMDLMEQVALRVQLOEQLRVQGEIDPKAOLLGCV 166
 DB 120 VQPYLDEFOKKWQEVERYRQKVEPLGALRESARQKLTLEQEKSLPLAEELR----- 172
 QY 167 DEAMALLQGLQSRVYVHTGRKELFHPYAESLVSGIGRNVOLHRSVADHPASPAPRLSR 226
 DB 173 DSARTHTVPLTKRLAPYSNELQORLAARLESIKKCGASLAEYQAKARKHLSVLSSEKARP 232
 QY 227 CVQVLSRKTLTKAKALHARIQONLDQLEELS 258
 DB 233 ALEDRLQGLIPVLESFKASVQNVLDHATKRLIN 264

RESULT 10
 ID APAL CHICK STANDARD: PRT: 264 AA.
 AC P08250:
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 GN APOA1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Archaeosuria; Aves; Chocornathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID:9031:
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88049703; PubMed=3118875;
 RA Byrnes L., Luo C.-C., Li W.-H., Yang C.-Y., Chan L.,
 RT "Chicken apolipoprotein A-I: cDNA sequence, tissue expression and
 RL evolution.";
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88152500; PubMed=3126099;
 RA Ferrarri S., Tarugi P., Drusiani E., Calandra S., Fregni M.,
 RT "The complete sequence of chick apolipoprotein AI mRNA and its
 RL expression in the developing chick.";
 RN 13
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8722301; PubMed=3108248;
 RA Rajavashisth T.B., Dawson P.A., Williams D.L., Shackelford J.E.,
 RT "Structure, evolution, and regulation of chicken apolipoprotein A-I.";
 RN 14
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92381402; PubMed=1512510;
 RA Lamon-Fava S., Sasstry R., Ferrarri S., Rajavashisth T.B.,
 RT "Evolutionary distinct mechanisms regulate apolipoprotein A-I gene
 RL expression: differences between avian and mammalian apoA-I gene
 J. Lipid Res. 33:831-842(1992)."

RN 15
 RP SEQUENCE OF 25-44.
 RX MEDLINE=83213468; PubMed=6406496;
 RA Shackelford J.E., Leberer H.G.,
 RT "Synthesis and secretion of apolipoprotein AI by chick breast
 muscle.";
 RN 16
 RP BIOLOGICAL CHEM. 258:7175-7180(1983).
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CC THE EFFLUX OF TISSUE CHOLESTEROL AND BY ACTING AS A COFACTOR FOR
 CC THE LECITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
 CC CHYLOMICRONS.
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOH FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).

DR EMBL: M17961; AAA48593.1;
 DR EMBL: M18746; AAA48594.1;
 DR EMBL: M25559; AAA48592.1;
 DR EMBL: M96012; AAA48597.1;
 DR PIR: S01453; LPECHAI.
 DR PIR: JH0471; JH0471.
 DR HSSP: P02647; LGM4.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein.
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 264
 FT DOMAIN 67 264
 FT REPEAT 67 88
 FT REPEAT 89 110
 FT REPEAT 111 121
 FT REPEAT 122 143
 FT REPEAT 144 165
 FT REPEAT 166 187
 FT REPEAT 188 209
 FT REPEAT 210 231
 FT REPEAT 232 242
 FT REPEAT 243 264
 FT CONFLICT 16 16
 FT CONFLICT 148 148
 SO SOURCE 264 AA: 30680 MW: 96832081E2AC5C2 CRC64:

Query Match 9.7%; Score 183; DB 1; Length 264;
 Best Local Similarity 21.1%; Pred. No. 0.00014;
 Matches 70; Conservative 62; Mismatches 113; Indels 86; Gaps 9;

QY 4 MAVLTWALALISAFSAIOARKCFW-----DYFSQT--SCDKGREVOIH 45
 DB 1 MCVLTWALVFL--TGSOARHFWQDEPRSSWDKIKDPATVYVDIVKSGREYVAQFEAS 57
 QY 46 KWARE-PATIKDLEQDINNMMKFELEKRLPLSGSEAPRLPODPVGMRRLOLELEEVKAR 104
 DB 58 SAAVKQDLDKALDNLDTLSAAAKLRDQMAVYKREVMKIDTALRAELTLDLEVK 117
 QY 105 ARLQPYAAEAHELVGNMIFGLKQOLKPYTMDLMEQVALRVQLOEQLRVQGEIDPKAOLLGCV 164
 DB 118 KIRIFLDFQSAKWEELRQKRLPLGALRESARQKLTLEQEKSLPLAEELR----- 177
 QY 165 DEAMALLQGLQSRVYVHTGRKELFHPYAESLVSGIGRNVOLHRSVADHPASPAPRLSR 224
 DB 178 -----HVEELRLNLABYSDLEKRL 197
 QY 225 SRCVQVLSRKTLTKAKALHARIQONLDQLEELSRAFAGTDEGAGDPQKMLSEEVROK 284

Db 198 SOKBEIRENGICPGASEYQAKVMQSLNLRKMT-----PLVGEFRER 240
 OY 285 LOAFRQDTYQIAAFTRAIDQETEEVQOOLA 315
 Db 241 LIPYAENLNKRLISF-----LDRLOKSA 264

RESULT 11

API2_ONCMY STANDARD: PRT: 262 AA.
 ID API2_ONCMY
 AC 057524;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Apolipoprotein A-I-2 precursor (APOA-I-2).
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Proteanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Shasta; TISSUE=Liver;
 RX MEDLINE=92235558; PubMed=1569377;
 RA Delcuve G.P., Sun J.M., Davie J.R.;
 RT "Expression of rainbow trout apolipoprotein A-I genes in liver and
 hepatocellular carcinoma."
 RL J. Lipid Res. 33:251-262(1992).

CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
 THE LECTIN CHOLESTEROL ACYLTRANSFERASE (LCAT) (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Extracellular (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE APOA / APOA4 / APOE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: AF042219; AAB96973.1;
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein.1.
 KM Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 23 POTENTIAL.
 FT CHAIN 24 262 APOLIPOPROTEIN A-I-2.
 FT DOMAIN 32 63 3 X APPROXIMATE TANDEM REPEATS.
 FT DOMAIN 64 262 10 X APPROXIMATE TANDEM REPEATS.
 FT REPEAT 64 85 1.
 FT REPEAT 87 107 2.
 FT REPEAT 108 118 3 (HALF-LENGTH).
 FT REPEAT 119 140 4.
 FT REPEAT 141 162 5.
 FT REPEAT 163 184 6.
 FT REPEAT 185 206 7.
 FT REPEAT 207 228 8.
 FT REPEAT 229 239 9 (HALF-LENGTH).
 FT REPEAT 240 262 10.
 SQ SEQUENCE 262 AA; 29680 MW; 234213D38B511A64 CRC64;

Query Match 9.58; Score 179; DB 1; Length 262;
 Best Local Similarity 22.58; Pred. No. 0.00024;
 Matches 60; Conservative 62; Mismatches 107; Indels 38; Gaps 7;

OY 8 LTNALLLSFSATQA-----RKGFMDYFSQSGKGRGEOHQKMAR-- 51
 Db 4 LALALITLLA-AATQAVPMQADAPSLQEHVAVAMETWAO-----VKETAQNSIDHLD 55

OY 52 --EPATLKDSLQEDLNNNKFL-----EKLRLPSGSEAPRLPDDPYGMRQLOEIEEYKA 105
 Db 56 DTEYEIKYQVLSQSLDNIQAQTAQESLAPSEAIQVLTATAVRAVEKADVEELRS 115
 OY 106 RLQPYMAEAEHVELGNLEGLROOLKPYTMDLMEQVALRVQEOLEQDRVVGEDTKQALLGG 165
 Db 116 QLEPRRAELKEVLDKHIDEYRKLEPLKIDIVEQRTLEAFRAVLEPVEELMRKAVSAN 175
 OY 166 VDEAMA----LIQIGQSRVYHHTGRFKEFLPFAESLVGIRHVOELHRSVAPHPASP 221
 Db 176 VEETRAKLMPVEYTRAKLTERLEELRLASPYAEYEQWAKAVGEVREKVVPLTDFK 235
 OY 222 ARLSRCVOYLSKLLTKAKALHARLDQ 248
 Db 236 GOLGPAAEQAKEL-----MALYETISQ 258

RESULT 12

APAL_BOVIN STANDARD: PRT: 265 AA.
 ID APAL_BOVIN
 AC P15497;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 OX [1]

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90348478; PubMed=2117227;
 RA O'Huigin C., Chan L., Li W.H.;
 RT "Cloning and sequencing of bovine apolipoprotein A-I cDNA and
 RT molecular evolution of apolipoproteins A-I and B-100."
 RL Mol. Biol. Evol. 7:327-339(1990).

RN [2]
 RP SEQUENCE OF 19-265.
 RX MEDLINE=92153895; PubMed=1739745;
 RA Sparrow D.A., Lee B.R., Laplaud M.P., Auboitron S., Beauchart D.,
 RA Chapman J.M., Goto A.M., Jr., Yang C.Y., Sparrow J.T.;
 RT "Plasma lipid transport in the preterm infant calf, Bos spp. primary
 RT structure of bovine apolipoprotein A-I."
 RL Biochim. Biophys. Acta 1123:145-150(1992).

RN [3]
 RP SEQUENCE OF 25-70.
 RX MEDLINE=90147795; PubMed=2105728;
 RA Auboitron S., Sparrow D.A., Beauchart D., Beauchart D., Sparrow J.T.,
 RA Laplaud M.P., Chapman J.M.;
 RT "Characterization and amino-terminal sequence of apolipoprotein AI
 RT from plasma high density lipoproteins in the preterm infant calf, Bos
 RT spp."
 RL Biochem. Biophys. Res. Commun. 166:833-839(1990).
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
 THE LECTIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
 CHYLOMICRONS.

CC -1- SIMILARITY: BELONGS TO THE APOA / APOA4 / APOE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: M35870; AAA30381.1; -

DR PIR: A34649; A34649.
 DR HSSP: P02647; 1AV1.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein; 1.
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 265
 FT DOMAIN 67 265
 FT REPEAT 67 88
 FT REPEAT 89 110
 FT REPEAT 111 121
 FT REPEAT 122 143
 FT REPEAT 144 165
 FT REPEAT 166 187
 FT REPEAT 188 209
 FT REPEAT 210 230
 FT REPEAT 231 241
 FT REPEAT 242 265
 FT CONFLICT 185
 FT SEQUENCE 265 AA; 30276 MW; 06A2681EA2ABA50F CRC64;
 Query Match 9.4%; Score 176; DB 1; Length 265;
 Best Local Similarity 22.0%; Pred. No. 0.00036;
 Matches 74; Conservative 58; Mismatches 103; Indels 102; Gaps 10;
 OY 5 AAVLTWALALLSAFSATQAR-----KGFMDFQSOTSGDKGR--VEQIHQ 47
 DB 3 AAVLTAVLVFL--TSSQARHFWMQDDPOSSMDRVDKDFATVYEAVIDSGRDYVAQEAS 59
 OY 48 KMARE-PATLKDSLEQDLDNNMKNFKLEKRLPSGSEARILPQDPVGMKROLOEELFEKKAR 106
 DB 60 ALGQNLKLLDMWDLTASTLSKREGLPVGTQEFMDNLEKETASLQEMNKDLEEVKQK 119
 OY 107 LQPYMAEAHELYGMNLGLRQOLKPYTMDMEQVALRVOELOEQLRVVSGEDTKAQLLGGV 166
 DB 120 VQPLDFQKKHVEVEIYQKVAPLCEPREGARQKVOELQDL----- 164
 OY 167 DEAMALLQGLQSRVYVHHTGFKELFHYAESLVSGIRHVOELHRSVAPAPASPARLSR 226
 DB 165 -----SPLAQLDRARAHYETLRQQLAPYS----- 190
 OY 227 CVOGLSKRLTIKAKAL-----HARIQONLDQRELSNAFACITGEGAGPPDM 276
 DB 191 --DDLRLRLARLEALKEGGSLAEYHAKASQDLKALGER-----AKP----- 231
 OY 277 LSEEVROLOAFRODTYLOIAFTRAIDQETEEVQOQ 313
 DB 232 VLEDLRQGLLPVLESLKVSIIA--AIDPASKKIMAQ 265
 RESULT 13
 APAL-CANFA STANDARD; PRT; 266 AA.
 AC P02648;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 CN APOA1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OC NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=90132271; PubMed=2515239;
 RA Luo C.-C., Li W.-H., Chan L.;
 RT "Structure and expression of dog apolipoprotein A-I, E, and C-I
 RT mRNAs: implications for the evolution and functional constraints of
 RT apolipoprotein structure";
 RL J. Lipid Res. 30:1735-1746(1989).

RN [2]
 RP SEQUENCE OF 25-266.
 RX MEDLINE=82142425; PubMed=6801039;
 RA Chung H., Randolph A., Reardon I., Heinrichson R.L.;
 RT "The covalent structure of apolipoprotein A-I from canine high
 RT density lipoproteins";
 RL J. Biol. Chem. 257:2961-2967(1982).
 RN [3]
 RP SEQUENCE OF 25-57 AND 262-265.
 RX MEDLINE=76210910; PubMed=179887;
 RA Nakai T., Whayne T.F., Tang J.;
 RT "The amino- and carboxyl-terminal sequences of canine apolipoprotein
 RT A-I";
 RL FEBS Lett. 64:409-411(1976).
 RN [4]
 RP SEQUENCE OF 25-37.
 RC TISSUE=Heart;
 RX MEDLINE=9816340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
 RT dog heart proteins";
 RL Electrophoresis 18:2795-2802(1997).
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CC THE LECTITHIN CHOLESTEROL ACYLTRANSFERASE (LCAT).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
 CC CHYLOMICRONS. SYNTHESIZED IN THE LIVER AND SMALL INTESTINE.
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
 DR PIR: A03092; LPDCA1.
 DR HSSP: P02647; 1AV1.
 DR HSC-2DPAGE; P02648; DOG.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein; 1.
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 266
 FT DOMAIN 67 266
 FT REPEAT 67 88
 FT REPEAT 89 110
 FT REPEAT 111 121
 FT REPEAT 122 143
 FT REPEAT 144 165
 FT REPEAT 166 187
 FT REPEAT 188 209
 FT REPEAT 210 231
 FT REPEAT 232 242
 FT REPEAT 243 266
 FT CONFLICT 168
 FT CONFLICT 202
 FT CONFLICT 235
 FT CONFLICT 264
 FT SEQUENCE 266 AA; 30196 MW; A3202620C28A869D CRC64;
 Query Match 9.4%; Score 175.5; DB 1; Length 266;
 Best Local Similarity 22.0%; Pred. No. 0.00039;
 Matches 72; Conservative 53; Mismatches 120; Indels 83; Gaps 8;
 OY 5 AAVLTWALALLSAFSATQARKFMD-----YFSQTSQDKGR--VEQIHQ 46
 DB 3 AAVLTAVLVFL--TSSQARHF--FMQDDPOSSMDRVDKDLATYVADAVKSGRDYVAQFEA 58
 OY 47 QKMARE-PATLKDSLEQDLDNNMKNFKLEKRLPSGSEARILPQDPVGMKROLOEELFEVKA 105
 DB 59 SALGQNLKLLDMWDLTASTLSKREGLPVGTQEFMDNLEKETASLQEMNKDLEEVKQ 118
 OY 106 RLQPYMAEAHELYGMNLGLRQOLKPYTMDMEQVALRVOELOEQLRVVSGEDTKAQLLGG 165
 DB 119 KVQPYLDDQKKHVEVEIYQKVAPLGSELNEGARKQLOEQL----- 164
 OY 166 VDEAMALLQGLQSRVYVHHTGFKELFHYAESLVSGIRHVOELHRSVAPAPASPARLS 225

DB 165 -----SPLAEELDRARITHVDALRAQALAPSDLLRERLA 198
 OY 226 RCYQVVISRKLTAKALHARIQONLDOLKEHLSRAFACTGTEGACHPDQMLSEEVHQR. 285
 DB 199 ARLEALEKGGASLAIEYHARASEQLSALGERKAPAL-----EDLRQGL 241
 OY 286 QAFRODTYLOIAAFTRAIDQTEEEVQOQ 313
 DB 242 LPVLESFVKVSLA---AIDEXTKLMAQ 266

RESULT 14
 APAL_ANAPL STANDARD; PRT: 264 AA.
 ID APAL ANAPL STANDARD; PRT: 264 AA.
 AC 04236;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 GN APOA1.
 OS Anas platyrhynchos (Domestic duck).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Anseriformes; Anallidae; Anas.
 OX NCBI_TaxID=8839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Pekin breed; TISSUE=Liver;
 RA Chen B., Lu X., Wang K., Xue H., Zeng W., Zhu D., Chi L.;
 RL Submitted (JAN1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CC THE LECTININ CHOLESTEROL ACYLTRANSFERASE (LCAT).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
 CC CHYLOMICRONS.
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U86131; AAB64381.1; -
 CC HSSP: P02647; IGM4.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein.
 KW Plasma; Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
 FT SIGNAL 1 18
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 264 APOLOPROTEIN A-I.
 FT DOMAIN 67 264 10 X APPROXIMATE TANDEM REPEATS.
 FT REPEAT 67 88 1.
 FT REPEAT 89 110 2.
 FT REPEAT 111 121 3 (HALF-LENGTH).
 FT REPEAT 122 143 4.
 FT REPEAT 144 165 5.
 FT REPEAT 166 187 6.
 FT REPEAT 188 209 7.
 FT REPEAT 210 231 8.
 FT REPEAT 232 242 9 (HALF-LENGTH).
 FT REPEAT 243 264 10.
 SQ SEQUENCE 264 AA; 30659 MW; 7546F4BAF2846900 CRC64;

Query Match 9.18; Score 171.5; DB 1; Length 264;
 Best Local Similarity 21.68; Prid. No. 0.00067;
 Matches 72; Conservative 63; Mismatches 100; Indels 99; Gaps 11;
 OY 7 VLTWALLLSAFSAIQARKGFWDYFSQTSQDKGRVEQIHQOKMARREPAITLKDLSQ----- 62

DB 5 VVTALLFL---TGTQARY-FWQH-DEQAPLDLRLVDVLYLTVAASGDALAQFEAS 59
 OY 63 -----DINNNKFLKRLPLSGSEAPRLPODPV-----GKKRQLEHLE 101
 DB 60 AVKQOLDL---KLADNLDL-GAAAKLREDMAPYKEVEEMMLKDTESLRAELTDLE 114
 OY 102 EVKARLQPYMAEAEHIEVGNWEGLRQOLKPYTMLEGVNVALRVQELQVAGEDTKAQ 161
 DB 115 EVKEKIRPFLDOFSAKWTLEQYRQRLAPAELEKTKOKVEMLOKLTVPVAFARDR 174
 OY 162 LLGGVDEAMALLQGLQSRVYHHTGRFELFIPYAESIVSGIRVQELHRSVAPAPASP 221
 DB 175 LRG-----HVEELRKLAPYSDEL 194
 OY 222 ARISRCQVVISRKLTAKALHARIQONLDOLKEHLSRAFACTGTEGACHPDQMLSEEV 281
 DB 195 QKLSQKLEIEIREKGIPOAEYQAKVVEQLSNLRKMT-----PLYODF 237
 OY 282 RQRLQAFRODTYLOIAAFTRAIDQTEEEVQOQ 315
 DB 238 KERLTPTVAENK-----TRFISL-LDELRQTVVA 264

RESULT 15
 APAL_PIG STANDARD; PRT: 265 AA.
 ID APAL_PIG STANDARD; PRT: 265 AA.
 AC P18648;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Apolipoprotein A-I precursor (Apo-AI).
 GN APOA1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93224154; PubMed=8468059;
 RA Bitchbauer A., Knipping G., Juritsch H., Aschauer H., Zechner R.;
 RT "Characterization of the apolipoprotein AI and CIII genes in the
 RT domestic pig";
 RL Genomics 15:643-652(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver, and Brain;
 RA Moeckel B., Zinke H., Flach R., Weiss B., Weller-Guetli H.,
 RT "Expression of apolipoprotein A-I in porcine brain endothelium in
 RT vitro";
 RL J. Neurochem. 62:788-798(1994).
 RN [3]
 RP SEQUENCE OF 34-265 FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=90132667; PubMed=2105375;
 RA Weller-Guetli H., Sommerfeldt M., Papandrikopoulou A., Mischek U.,
 RA Bonitz D., Frey A., Grupe M., Scheerer J., Gassen H.G.;
 RT "Synthesis of apolipoprotein A-I in pig brain microvascular
 RT endothelial cells";
 RL J. Neurochem. 54:444-450(1990).
 RN [4]
 RP SEQUENCE OF 105-265 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93154581; PubMed=8428656;
 RA Tieu V.N., Hasler-Rapacz J., Rapacz J., Black D.D.;
 RT "Sequences and expression of the porcine apolipoprotein A-I and C-III
 RT mRNAs";
 RL Gene 123:173-179(1993).
 RN [5]
 RP SEQUENCE OF 25-265.
 RA Hasler-Rapacz J.O., Chaudhary R., Chowdhary B.P., Tieu V.N.,

RA Jackson K., Gustavsson I., Rapacz J.;
 RI Submitted (Oct-1995) to the SWISS-PROT data bank.
 RN [6]
 RP SEQUENCE OF 25-34.
 RX MEDLINE-76184721; PubMed-178359;
 RA Mahley R.W., Weisgraber K.H., Innerarity T., Brewer H.B. Jr.;
 RT "Characterization of the plasma lipoproteins and apoproteins of the
 RT Erythrocytes of the monkey."
 RT Biochemistry 15:1928-1933(1976).
 CC -1- FUNCTION: APOA-1 PARTICIPATES IN THE REVERSE TRANSPORT OF
 CC CHOLESTEROL FROM TISSUES TO THE LIVER FOR EXCRETION BY PROMOTING
 CC CHOLESTEROL EFFLUX FROM TISSUES AND BY ACTING AS A COFACTOR FOR
 CC THE LECTININ CHOLESTEROL ACYLTRANSFERASE (LCAT).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAJOR PROTEIN OF PLASMA HDL, ALSO FOUND IN
 CC CHYLOMICRONS. SYNTHESIZED PREDOMINANTLY IN THE INTESTINE AND THE
 CC LIVER.
 CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: L00626; AAA30992.1; -
 DR EMBL: X69477; CAA49234.1; -
 DR EMBL: X17057; -; NOT_ANNOTATED_CDS.
 DR EMBL: X59414; CAA42050.1; -
 DR PIR: J00704; J00704.
 DR PIR: A05311; A05311.
 DR PIR: A46018; A46018.
 DR PIR: S21830; S21830.
 DR PIR: S31394; S31394.
 DR HSSP: P02647; IAVI.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein; 1.
 DR KJ Plasma: Lipid transport; HDL; Cholesterol metabolism; Repeat; Signal.
 FT SIGNAL 1 18
 FT PROPR 19 24
 FT CHAIN 25 265
 FT DOMAIN 67 265
 FT REPEAT 67 88
 FT REPEAT 89 110
 FT REPEAT 111 121
 FT REPEAT 122 142
 FT REPEAT 144 165
 FT REPEAT 166 187
 FT REPEAT 188 209
 FT REPEAT 210 230
 FT REPEAT 231 241
 FT REPEAT 242 265
 FT CONFLICT 108 108
 FT CONFLICT 143 143
 FT CONFLICT 173 173
 FT CONFLICT 180 180
 FT CONFLICT 185 186
 FT CONFLICT 209 209
 FT CONFLICT 224 224
 SO SEQUENCE 265 AA; 30325 MM; 2C6E578318ECF69C CRC64;

Query Match 9.1%; Score 170; DB 1; Length 265;
 Best Local Similarity 22.3%; Pred. No. 0.00083;
 Matches 73; Conservativity 52; Mismatches 120; Indels 82; Gaps 8;

OY 5 AAVLTWALALISAFSATOAR-----KGFWDYFSQNSGDKR--VVGQIHQO 47
 DB 3 AVVLTWALVFL--TGSQARHFWQDDPOSMDRYKDFATYVDAIKDSGRDYVAQFEAS 59
 OY 48 KVARB-PATLKDSLQDQNNMKNKFLKLRPLSGSEAPRLQDPVGMKROLOEELEVKAR 106

DB 60 ALGKHNLMLKLDNWDLSGSTFTKVRQQLGPTQGFMDNLEKFEALRQEMSKILFEYKKK 119
 OY 107 LOPYMAEAEHELVCWNLBGLRQQLKPYTMDLMEOVALLRVQELQRLVYGEDTAQLLGCV 166
 DB 120 VQYLDLDFONKMOEEMETVRQKMAPLGAEREGARQKVCLOEKL----- 164
 OY 167 DEAMALLQGLQSVVHHTGCFKELFHPYAESIVSGICRHVQELHRSVAPHPAPASPARISR 226
 DB 165 -----SFLAEELDRLEAHEALRQNVAPYSDDLROKMA 199
 OY 227 CVOVLSRKLLTKAKALHARIOQLDQRLRELSRAFGTGEAGDPQMLSEEVQRQLQ 286
 DB 200 RPEALKEGCSLAE-VQAKAQEQLKALCFKAKVAL-----KDLQGLL 241
 OY 287 AFRODTYLOIATFRAIDQTEVEYQOQ 313
 DB 242 PVLENLKVSTLA--AIDEASKRLINQ 265

Search completed: January 23, 2003, 17:44:23
 Job time: 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2003, 17:36:06 : Search time 76 seconds
(without alignments)
992.280 Million cell updates/sec

Title: US-09-842-364-3
Perfect score: 1877
Sequence: 1 MASMAVLTMALALSAFSA.....LMEDITHSLHDGSHLGR 366

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1864	99.3	363	4 Q9UBJ3	Q9UBJ3 homo sapien
2	1310	69.8	368	11 Q99P64	Q99P64 mus musculu
3	1301	69.3	368	11 Q91X90	Q91X90 mus musculu
4	1293	68.9	367	11 Q90UH3	Q90UH3 rattus norv
5	316	16.8	366	13 Q93601	Q93601 gallus gall
6	302.5	16.1	395	11 Q91XF8	Q91XF8 mus musculu
7	299.5	16.0	395	11 Q9DBN0	Q9DBN0 mus musculu
8	289.5	15.4	435	11 Q01488	Q01488 mus musculu
9	188	10.0	244	4 Q13784	Q13784 homo sapien
10	172	9.2	264	11 Q922L4	Q922L4 mesocricetu
11	165	8.8	990	4 Q9NYF9	Q9NYF9 homo sapien
12	164	8.7	317	6 Q9GLM7	Q9GLM7 pongo pygma
13	163.5	8.7	162	16 Q91316	Q91316 pseudomona
14	163	8.7	259	13 Q98TGI	Q98TGI anguilla ja
15	163	8.7	317	6 Q9GLM8	Q9GLM8 gorilla gor
16	162	8.6	313	6 Q9GLC0	Q9GLC0 tupia glis

17	162	8.6	317	6 Q9GJ03	Q9GJ03 pan troglod
18	160	8.5	317	6 Q9GLM6	Q9GLM6 hylobates 1
19	159.5	8.5	275	13 Q9PT02	Q9PT02 oncomychnu
20	156.5	8.3	259	13 Q98TGS	Q98TGS anguilla ja
21	155.5	8.3	263	11 Q08855	Q08855 mus musculu
22	155.5	8.3	2310	5 Q9NKR1	Q9NKR1 leishmania
23	153.5	8.2	259	13 Q98TGA	Q98TGA anguilla ja
24	152.5	8.1	258	11 Q08877	Q08877 rattus norv
25	152.5	8.1	263	11 Q09042	Q09042 mus musculu
26	151.5	8.1	1956	5 Q20641	Q20641 caenorhabdi
27	151	8.0	241	6 Q9TS49	Q9TS49 erinaceus e
28	151	8.0	466	4 Q60374	Q60374 homo sapien
29	151	8.0	1313	4 Q75033	Q75033 homo sapien
30	151	8.0	2354	5 Q9NKT9	Q9NKT9 leishmania
31	150.5	8.0	591	5 Q17117	Q17117 caenorhabdi
32	149.5	8.0	258	11 Q09054	Q09054 rattus norv
33	147.5	7.9	1248	16 Q8U101	Q8U101 agrobacteri
34	147	7.8	800	3 Q96X03	Q96X03 emericeila
35	144.5	7.7	259	13 Q98TGS	Q98TGS anguilla ja
36	144.5	7.7	259	13 Q98TGS	Q98TGS anguilla ja
37	143.5	7.6	263	13 Q98TGS	Q98TGS anguilla ja
38	141.5	7.5	1000	2 Q54762	Q54762 synecococc
39	140	7.5	2779	5 Q9W4N7	Q9W4N7 drosophila
40	139.5	7.4	590	6 Q9BGY1	Q9BGY1 macaca fasc
41	139.5	7.4	806	11 Q8VD04	Q8VD04 mus musculu
42	139	7.4	174	13 Q98S13	Q98S13 cyprinus ca
43	138.5	7.4	552	5 Q967G6	Q967G6 trichinella
44	138.5	7.4	1208	4 Q9P2M7	Q9P2M7 homo sapien
45	137.5	7.3	1210	11 Q91VW5	Q91VW5 mus musculu

ALIGNMENTS

RESULT 1

ID	Q9UBJ3	PRELIMINARY:	PRT:	363 AA.
AC	Q9UBJ3:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	Regeneration associated protein 3.			
GN	RAP3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;			
KA	Van der Vliet H.N., Keltzma P.H., Boers W., Chamuleau R.A.P.M.;			
RT	*Identification of the human analog of rat RAP3.*;			
RI	Submitted (NOV-1999) to the EMBL/GenBank/DDJ databases.			
DR	EMBL: AF202890; AAF25662.1; -			
DR	EMBL: AF202899; AAF25661.1; -			
DR	InterPro: IPR000074; Apolipoprotein.			
DR	Pfam: pf01442; Apolipoprotein; 1.			
DR	SEQUENCE 363 AA; 40923 MW; 9E09CE97205D8A7D CRC64;			
QY	4 MAAVLTNALALISAFSAQAARKGFWDFSGTSGDKGRVEQIHQKMAREPATLKSLEQD 63	Query Match	99.3%	Score 1864; DB 4; Length 363;
QY	1 MAAVLTNALALISAFSAQAARKGFWDFSGTSGDKGRVEQIHQKMAREPATLKSLEQD 60	Best Local Similarity	100.0%	Pred. No. 1.8e-112;
QY	64 LNNNNKFLKRLRPGSSSEAPRLPDDPYGMKROLEELLEVYKARQIPMAAHLEVGNLE 123	Matches 363; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	61 LNNNNKFLKRLRPGSSSEAPRLPDDPYGMKROLEELLEVYKARQIPMAAHLEVGNLE 120			
QY	124 GLRQOLKPYTMDLMEQVALRQVLEQLOELRVVGEPTKQQLGVDENALLOGLSRVNH 183			

Db 121 GLRQQLKPYTMDLMEQVALRVGELOQLRVGEDTKAQLLGGVDKAMALLQGLQSRVNH 180
 QY 184 TGRFELFHPYAESLVSGIGHVQVLEHRSVAPAPASPARLSRCVOVLSKRLTKAKALH 243
 Db 181 TGRFELFHPYAESLVSGIGHVQVLEHRSVAPAPASPARLSRCVOVLSKRLTKAKALH 240
 QY 244 ARIQONLQRLRELSRAFGTGTGAGGPDPMLEEVQRLOAFRODYLQIAAFTRAI 303
 Db 241 ARIQONLQRLRELSRAFGTGTGAGGPDPMLEEVQRLOAFRODYLQIAAFTRAI 300
 QY 304 DQETEEVOOQLAPPPGHSAPAFEPQOTDSKVLSKLARLDLMDLMDITHSLHDOGHSHL 363
 Db 301 DQETEEVOOQLAPPPGHSAPAFEPQOTDSKVLSKLARLDLMDLMDITHSLHDOGHSHL 360
 QY 364 GDP 366
 Db 361 GDP 363

RESULT 2

099P64 ID 099P64 PRELIMINARY: PRT: 368 AA.
 AC 099P64
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Regeneration-associated protein 3.
 CN APOA5 OR RAP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J TISSUE=LIVER.
 RA Van der Vliet H.N., Groenink M., Leegwater A.C.J., Chamuleau R.A.F.M.;
 RT Identification of the mouse analog of rat RAP3.
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF337059; AAC49600.1; -
 DR MGD: MGI:1913363; APOA5.
 DR InterPro: IPR000074; APOlipoprotein.
 DR Pfam: PF01442; APOlipoprotein; 1.
 SQ SEQUENCE 368 AA; 41320 MW; A55DF15BD0784696 CRC64;

Query Match 69.8%; Score 1310; DB 11; Length 368;
 Best Local Similarity 72.6%; Pred. No. 8.7e-77;
 Matches 267; Conservative 28; Mismatches 63; Indels 10; Gaps 5;

QY 4 MAALVTMALALLSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSLEOD 63
 Db 1 MAALVTMALALLSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSLEOD 56
 QY 64 LNNMKFLEKLRLPL--SGSEAPRLPDDPGMRKQLOEELLEVYKARLOPYMAEHLVGN 121
 Db 57 LNNMKFLEKLRLPL--SGSEAPRLPDDPGMRKQLOEELLEVYKARLOPYMAEHLVGN 116
 QY 122 LEGLRQQLKPYTMDLMEQVALRVGELOQLRVGEDTKAQLLGGVDKAMALLQGLQSRV 181
 Db 117 LEGLRQQLKPYTMDLMEQVALRVGELOQLRVGEDTKAQLLGGVDKAMALLQGLQSRV 176
 QY 182 HHTGFKELFHPYAESLVSGIGHVQVLEHRSVAPAPASPARLSRCVOVLSKRLTKAKA 241
 Db 177 HHTGFKELFHPYAESLVSGIGHVQVLEHRSVAPAPASPARLSRCVOVLSKRLTKAKA 236
 QY 242 LHAHQONLQRLRELSRAFGTGTGAGGPDPMLEEVQRLOAFRODYLQIAAFTRAI 298
 Db 237 LHAHQONLQRLRELSRAFGTGTGAGGPDPMLEEVQRLOAFRODYLQIAAFTRAI 295
 QY 299 FTRALDQETEEVOOQLAPPPGHSAPAFEPQOTDSKVLSKLARLDLMDLMDITHSLHDO 358
 Db 296 FTRALDQETEEVOOQLAPPPGHSAPAFEPQOTDSKVLSKLARLDLMDLMDITHSLHDO 355
 QY 359 GHSGLSDP 366

Db 356 GHSGLSDP 363

RESULT 3

091X90 ID 091X90 PRELIMINARY: PRT: 368 AA.
 AC 091X90
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Unknown (protein for MGI:18812).
 GN APOA5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER.
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC011198; AAH1198.1; -
 DR MGD: MGI:1913363; APOA5.
 DR InterPro: IPR000074; APOlipoprotein.
 DR Pfam: PF01442; APOlipoprotein; 1.
 SQ SEQUENCE 368 AA; 41237 MW; 174DB2481B88A660 CRC64;

Query Match 69.3%; Score 1301; DB 11; Length 368;
 Best Local Similarity 72.0%; Pred. No. 3.3e-76;
 Matches 265; Conservative 29; Mismatches 64; Indels 10; Gaps 5;

QY 4 MAALVTMALALLSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSLEOD 63
 Db 1 MAALVTMALALLSAFSAATQARKGFWDYFSQTSQDKGRVQIHOQKMARPPATLKDSLEOD 56
 QY 64 LNNMKFLEKLRLPL--SGSEAPRLPDDPGMRKQLOEELLEVYKARLOPYMAEHLVGN 121
 Db 57 LNNMKFLEKLRLPL--SGSEAPRLPDDPGMRKQLOEELLEVYKARLOPYMAEHLVGN 116
 QY 122 LEGLRQQLKPYTMDLMEQVALRVGELOQLRVGEDTKAQLLGGVDKAMALLQGLQSRV 181
 Db 117 LEGLRQQLKPYTMDLMEQVALRVGELOQLRVGEDTKAQLLGGVDKAMALLQGLQSRV 176
 QY 182 HHTGFKELFHPYAESLVSGIGHVQVLEHRSVAPAPASPARLSRCVOVLSKRLTKAKA 241
 Db 177 HHTGFKELFHPYAESLVSGIGHVQVLEHRSVAPAPASPARLSRCVOVLSKRLTKAKA 236
 QY 242 LHAHQONLQRLRELSRAFGTGTGAGGPDPMLEEVQRLOAFRODYLQIAAFTRAI 298
 Db 237 LHAHQONLQRLRELSRAFGTGTGAGGPDPMLEEVQRLOAFRODYLQIAAFTRAI 295
 QY 299 FTRALDQETEEVOOQLAPPPGHSAPAFEPQOTDSKVLSKLARLDLMDLMDITHSLHDO 358
 Db 296 FTRALDQETEEVOOQLAPPPGHSAPAFEPQOTDSKVLSKLARLDLMDLMDITHSLHDO 355
 QY 359 GHSGLSDP 366
 Db 356 GHSGLSDP 363

RESULT 4

090UH3 ID 090UH3 PRELIMINARY: PRT: 367 AA.
 AC 090UH3
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Regeneration associated protein 3.
 GN RAP3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

Db 185 NLKKFNFNMEELGHTLTPRANEKATIDONLEDRSLAPLYTVQVEKLNHMEGLAQ 244

QY 235 I,TLKAKALHARIQONLQ,RELSRAQF-----GREBACGDPOMLSVEVQRIDAR 289

Db 245 MKKNAEIQTVTSAKIDOLQNLAP,LEDVQSKVKGNTBGLQKSLIEDLNRLQOEVEFR 304

QY 290 QDTYIQIAFLRAIDQETEEVQOOLAPPPGHSAPFAEPQOTDSKV---LSLQARLDD 346

Db 305 KTVPEPMCEMFKALVQOL,EDFRQOLCP-----NSGEVESHLSLEKSLRE 349

QY 347 LMEDITTHSDQG 359

Db 350 KVNSEMFSTLEKKG 362

Db	125	GENNOKIDENHKRAVAVDQDQINFQTOGEMKQDLPTRYQIBQOTTIKEVNDLHNSMPLAT	184
QY	175	GIQSRVNHHTGRFELHPRYAESVSGIRHVOELHNSVAPHAPASPARISRCVQVUSRK	234
Db	185	NLKDFNNMMEELKGHLTPRANELKATIDQMLPLKRSJLAPTVQVOEKLNHQMEGJAC	244
QY	235	LTLKKALKHAKIQOQMLDQREELSRAFGT-----GTEBAGQDPQMLSEEVQORLOAF	289
Db	245	MKNMNEELQTVSAKIDLOQKNLAPLVEDVQSKVKGCTPQLOKSLIEDLNQOLEQOEEFR	304
QY	250	QDTYQIQAFLRAIDQELERHQOQOLARRPRHNSAFARHPQOOTSCKY--LSKIQALDD	346
Db	305	RVEBEMGEMFKKALVQOLEQDQDGR-----NSGEVSHSLPLEKESLE	349

RESULT 7			
Q9DBNO	PRELIMINARY:	PT:	395 AA.
ID	Q9DBNO:		
AC	Q9DBNO:		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Adult male liver cDNA, RIKEN full-length enriched library, clone:1300002K10, full insert sequence.		
GN	AP0A4.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=LIVER;		
RC	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Komu H., Adachi J., Fukuda S.,		
RA	Alawa K., Iwawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,		
RA	Salto T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King R., Kochia H.,		
RA	Kuchl P., Lewis S., Matsuo T., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schmitt L.M., Staudil F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Botfield D., Bojunga N., Carinici P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Guisticich S., Hill D., Hoffman M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni F., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitlaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuki S.,		
RA	Hayashizaki Y.;		
RT	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
DR	EMBL: AK004856; BAB23620.1; -		
DR	MED: MGI:88051; APOA4.		
DR	InterPro: IPR000074; Apolipoprotein.		
DR	Pfam: PF01442; Apolipoprotein. 2.		
DR	SEQUENCE 395 AA; 45044 MW; 4102D84ACBD0182A CRC64;		

```

DB          350  KVN5FMSTLTKKKG 362          :| :|
RESULT 8
ID 001488      PRELIMINARY;          PRT;      435 AA.
AC 001488;
DT 01-NOV-1996 (TREMblrel, 01, Created)
DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Variant apolipoprotein A-IV precursor (APOA-IV).
OS APOA-4.
OS Mus musculus castaneus (southeastern Asian house mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91286309; PubMed=1648102;
RA Reue K., Leete T.H.;
RT "Genetic variation in mouse apolipoprotein A-IV due to insertion and
RL deletion in a region of tandem repeats.";
RL J. Biol. Chem. 266:12715-12721(1991).
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIPASE BY APOC-II; POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS. IT MAY PLAY A ROLE IN THE
CC INTRAVASCULAR METABOLISM OF HDL. POTENTIAL ROLE IN CELLULAR
CC CHOLESTEROL EFFLUX.
CC -1- TISSUE SPECIFICITY: INTESTINE, LIVER, AND PLASMA.
CC -1- MISCELLANEOUS: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS
CC (EACH 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A & B, RELATED 11
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA
CC HELICAL, & MANY OF THESE HELICES ARE AMPHIPATHIC. THEY MAY
CC THEREFORE SERVE AS LIPID - BINDING DOMAINS WITH
CC LECITHIN:CHOLESTEROL ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- MISCELLANEOUS: THE APOA-IV POLYMORPHISMS ARE CLASSIFIED BASED ON
CC THE OCCURENCE OF THREE, FOUR, OR FIVE E-Q-A-Q/ E-Q-Q-Q/ E-Q-V-Q
CC REPEAT UNITS.
DR EMBL: M64250; AAA37216.1; -.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 2.

```

[illegible]

KW	Plasma: multigene family; Polymorphism: cholesterol metabolism.			
FT	SIGNAL	1	20	POTENTIAL.
FT	CHAIN	21	435	VARIANT APOILOPROTEIN A-IV.
FT	DOMAIN	374	393	TANDEN REPEATS.
FT	REPEAT	374	377	
FT	REPEAT	378	381	1.
FT	REPEAT	382	385	2.
FT	REPEAT	386	389	3.
FT	REPEAT	390	393	4.
FT	REPEAT	390	393	5.
SO	SEQUENCE	435 AA;	49254 MM;	B2AF55E8AE456B2 CRC64;
	Query Match	15.48;	Score 289.5;	DB 11; Length 435;
	Best Local Similarity	23.98;	Pred. No. 5, 8e-11;	
	Matches 89; Conservative	85;	Mismatches 166;	Indels 33; Gaps

```

QY 5 AAVLTWALLLSA----FSATQARKGFMDYFSQTSQD-KGRVEIHQOKKAREPATL-KD 58
DB 5 AAVLTWALLVAITQTRAKVTSDQVANYVMDFTLHLSNNAKELAQOFQRTDTQULSTLFD 64
QY 59 SLEODLNKKLEKLEKRLPSGSEAPRLPODPVGRROLOEELEEVKARLOPYMAEAHEL 118
DB 65 KLGDASTYADVGNKLVLPFYVQSLGSLAKETEVKEIKELLEDLDLDRMHPHANKVTQTF 124
QY 119 GNNLEGLROOLKPYTMDLMEQVALRYOELOEOLRVYGEEDTKAOLIGVD---EAMALLQ 174
DB 125 GEMMOKLOELKELKPYAVDLODQINTQTOEMKLOLTPYIQORQOTIKEDVILHTSMPLAT 184
QY 175 GLOSRVYHHTGKREKLEFHPAESLVSIGRHYOELHRSVAPHAPASPARLSRCVOYLSRK 234
DB 185 NIKDKFRNNKEELKGLHTPHANPLKATTDONLDRKSLAPLTVCQVQKLNHOHGLAQ 244
QY 235 LTLKAKALHARIQONLDQREELSRAFAGT---GTEGAGPDPOMLSEEVQRLQAFR 289
DB 245 MKKNABELOTQVSAKIDQLOKKNLAPLVEDQSKVKNTEGLQSKLDLNRQLODQVEEFR 304
QY 290 QDTYLOLAFTRALIDQTEHVOQOLAPPPGHSAFAPRYOOTHISGV---LSKIQARLDD 346
DB 305 RTVEPGEEMFNKALVOQLQEFROQLCP-----NSGEVSHLSLEKSLRE 349
QY 347 LMEDITHSLIDQG 359
DB 350 KVSFMSITLEKKG 362

```

RESULT 9
013784 PRELIMINARY: PRT: 244 AA.

```

ID 013784:
AC 013784:
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DR APOA4 protein (Fragment).
GN APOA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX TISSUE=LIVER;
RC MEDLINE=86016704; PubMed=3931073;
RA Karathanasis S.K.;
RT "Apolipoprotein multigene family: tandem organization of human
RT apolipoprotein AI, CII, and AIV genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6374-6378(1985).
DR EMBL: M10373; AAB59516.1;
DR HSSP: P02649; 1B24.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 1.
FT NON_TER 1
SQ SEQUENCE 244 AA; 28157 MW; 01616ACB53F12734 CRC64;

```

Query Match 10.0%; Score 188; DB 4; Length 244;
Best Local Similarity 26.1%; Pred. No. 9.7e-05;
Matches 59; Conservative 44; Mismatches 97; Indels 26; Gaps 4;

```

QY 107 LQPYMAEAHELVGWNLGEG--LRQOLKPYTMDLMEQVALRYOELOEOLRVYGEEDTKAOLIGV 166
DB 1 LEPYVAQQLRTQVTVQAEQRLQRLDPLAQRMHRYLRENNASLSQSLRPHADELAKKIDQV 60
QY 167 DEAMALLQGLQSRVYHHTGKREKLEFHPAESLVSIGRHYOELHRSVAPHAPASPARLSR 226
DB 61 EE-----LKGRLT-----PYADEFKVKIDQTEVIELRSLAPYAODTOEKLNH 102
QY 227 CVOVLSRKLTAKAKALHARIQONLDQREELSRAFAGT---GTEGAGPDPOMLSEEV 281
DB 103 QLEGITFQMKKNABEELKARISASAEELRORLAPLAEVDVGNLKGNGTEGLQSLAEIGHL 162

```

```

QY 282 RQRLQAFRODTYIQLAFAFTRALIDQTEHVOQOLAP---PPGHSAF 324
DB 163 DQOVEFRFRREVPYGENFNKALVOQMEOLKOLGPHAGDVEGHSF 208

```

RESULT 10

```

ID 092214 PRELIMINARY: PRT: 264 AA.
AC 092214:
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DR Apolipoprotein A-I.
GN APOA1.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OX NCBI_TaxID=10036;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=GOLDEN SYRIAN; TISSUE=INTESTINE;
RX MEDLINE=99061559; PubMed=9843713;
RA Wu J.Y., Reeves S.K., Wang Y.R., Wu Y., Lei P.P., Lei K.Y.;
RT "Zinc deficiency decreases plasma level and hepatic mRNA abundance of
RT apolipoprotein A-I in rats and hamsters.";
KL Am. J. Physiol. 275:C1516-C1525(1998).
DK EMBL: AF046919; AAC98484.1;
DR HSSP: P02647; IAV1.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 1.
KW Lipoprotein.
SQ SEQUENCE 264 AA; 30739 MW; 280B22F4C0B0129 CRC64;

```

Query Match 9.2%; Score 172; DB 11; Length 264;
Best Local Similarity 21.8%; Pred. No. 0.0011;
Matches 61; Conservative 57; Mismatches 112; Indels 50; Gaps 8;

```

QY 9 TMAIALISAF-SATQAR-----KGFMDYFSQTSQDQGR--VEQIHQOKKA 50
DB 3 TVLVAVALVELTSGQARHFWQDDPQTPMDRVKDFATVYVDAAVKGDSREVYSQFETSLG 62
QY 51 RE-PATLKDSLEODLNKKLEKRLPSGSEAPRLPODPVGRROLOEELEEVKARLOP 109
DB 63 KQNLINLLEWMDLPGSVGLOEOLGCPVTOEFMDNLKETEEMLRRENNKLEEVKAYOP 122
QY 110 YMAEAHELVGWNLGEG--LRQOLKPYTMDLMEQVALRYOELOEOLRVYGEEDTKAOLIGV 167
DB 123 YLDQFO--TKWQEEVALYRQKMEPLGALRDGAKRQKLOELQELKLTPLGED----- 170
QY 168 KAMALLQGLQSRVYHHTGKREKLEFHPAESLVSIGRHYOEL-----HRSVAPHAP 218
DB 171 -----LDRMRHNVDAALRTKMTPSDQMRDLERLQALDSDPTLAELYHTKADHLK 222
QY 219 ASPARLSRCVOYLSRKLTAKAKALHARIQONLDQREELS 258
DB 223 AFGEKAKPALEDLRLQGLMPFESFKTRIMSMVDEASKLN 262

```

RESULT 11

```

ID 09NYF9 PRELIMINARY: PRT: 990 AA.
AC 09NYF9:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DR Golgi matrix protein GM130.
GN GOLGA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PLACENTA:
 RA Weide T.B., Bayer M., Barnekow A.:
 RT "Human golgi matrix protein 130, hucM130, a protein highly homologous
 to GM130 from rat."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF248953; AAF65550.1;
 KW Matrix protein.
 SQ SEQUENCE 990 AA; 111658 MW; 2380FBED11B8DAF5 CRC64;

Query Match 8.8%; Score 165; DB 4; Length 990;
 Best Local Similarity 22.0%; Pred. No. 0.017;
 Matches 91; Conservative 63; Mismatches 128; Indels 132; Gaps 16;

QY 15 LSASFATARKGFMD-----YSSQTSQDKGRVQIHOQKAREPATLKDSL 60
 DB 260 LSASFATARKGFMD-----YSSQTSQDKGRVQIHOQKAREPATLKDSL 319
 QY 61 LODLNMMKFLKRLPL-----SCSEAPRLPODPVGMRRQIQEILEYKARIQPYMAEAH 115
 DB 320 QLNLEELQKLEMTLELLQGFSSRCEAPDANO-----QLQAMNEF-RAQLEAHLGGVM 371
 QY 116 ELVGNLEGLRQOLKPYMDIMEQVAL---RVQELQEOI.RVVGEDTKAQLLGVDEAMAL 172
 DB 372 E---SVRLQEMENDKVAENLKGSAMWROHMOQSEQVHTLRKEKCSM----- 417
 QY 173 LOGIQRVNHHTGRKELFHPYAEISVSGIRHVOELHRSVA-----PHAPASP 221
 DB 418 -----SRVOELTSLAELRNQMAEPPEPPACP 446
 QY 222 ARLSKCVQ---LSRKLTLAKALHANIQON---IDQIRE---ELSRFACTGTEE 268
 DB 447 SEVEOLOAEAEHLEKELGAGLOAQOVQDNESRLNRCOEELTLERAAELMGQ- 505
 QY 269 GAGDPQOMLSEVGRQIQAQPDYTIQIAAFTRAIDQETEVQOOLAPPPHSAFAPEF 328
 DB 506 -----AARROILTTMONDR---TTISRALSO-NREIKEDLAELOSGQVFKLTMEN 551
 QY 329 QQRTDS-----GKVLSTKIOARLDLMDI-----THSLHDQGHSHLG 364
 DB 552 MEITSALQSEQHVKEIKGKIGLOEKISLKEVELKSGVAGSIIQQQRDQYIG 605

RESULT 12
 Q9GLM7 PRELIMINARY; PRT; 317 AA.
 AC O9GLM7;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Apolipoprotein E.
 GN APOE.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kogaev E.I., Dvorianchikov G.A., Riazanskaja N.N.:
 RT "APOE gene evolution in Homnidae."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF200505; AAG28580.1;
 DR EMBL: AF200503; AAG28580.1; JOINED.
 DR EMBL: AF200504; AAG28580.1; JOINED.
 DR HSP; P02649; JOEF.
 DR InterPro: IPR000074; Apolipoprotein.
 DR Pfam: PF01442; Apolipoprotein; 1.
 KW Lipoprotein.
 SQ SEQUENCE 317 AA; 36094 MW; 6B08CE5B033BC2D CRC64;

Query Match 8.7%; Score 164; DB 6; Length 317;
 Best Local Similarity 21.4%; Pred. No. 0.0047;

Matches 67; Conservative 58; Mismatches 120; Indels 68; Gaps 8;

QY 10 WALALLSASFATQARKGFMDY--FSQTSQDKGRVQIHOQKAREPATLKDSL 67
 DB 44 WELAL-----GKWDYLRWQVLTSEQVO--ELLSSQVOTELTALMDETKELKAY 92
 QY 68 NKFL-EKLRPLSGSAPRLPODPVGMRRQIQEILEYKARIQPYMAEAHGVNMLDGLR 126
 DB 93 KSELEQLPYAEETRALSLKELQAARLQADMDVGRVQYRGEVQAMGOSTEELR 152
 QY 127 QOLKEPYTMDLMEQVALRVQELQEOI.RVVGEDTKAQLLGVDEAMALLOGIQRVNHHTGR 186
 DB 153 ARLASHLKLKRLRLRDADDLQKRLAAYQACARGAENG-----SALIKERLIG 201
 QY 187 FKELFHPYAEISVSGIRHVOELHRSVAPIAPASPARLSRCVOYLKRLTLAKALHARI 246
 DB 202 LVEQGRVRAATYGVSAVGRPIQER-----AQAWGERLRARMEKMGSR 243
 QY 247 QONILQOLREELSRAPAGCTERGACPDHOMLSEVGRQIQAQPDYTIQIAAFTRAI--- 303
 DB 244 RDRIDEVEQVA-----EVRALKEQADQIIRQAEAFQARLKS 282
 QY 304 -DQETEVQOOLA 315
 DB 283 FEPLVEDMOROWA 295

RESULT 13

Q9I316 PRELIMINARY; PRT; 1162 AA.
 ID Q9I316;
 AC Q9I316;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein PA1527.
 GN PA1527.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-ATCC 15692 / PA01;
 RA MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzer L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Keizer J., Sailer M.H., Hancock R.F.W., Lory S., Olson M.V.:
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AE004581; AAG04916.1;
 DR InterPro: IPR003439; ABC transport.
 DR InterPro: IPR003405; SMC_C.
 DR InterPro: IPR003395; SMC_N.
 DR Pfam: PF02483; SMC_C; 1.
 DR Pfam: PF02463; SMC_N; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1162 AA; 131580 MW; A62DDCC73015ACB5 CRC64;

Query Match 8.7%; Score 163.5; DB 16; Length 1162;
 Best Local Similarity 21.8%; Pred. No. 0.026;
 Matches 105; Conservative 59; Mismatches 153; Indels 165; Gaps 20;

QY 3 SMAALVFWALLLSASFATQARKGFMDYFSQTSQDKGRVQIHOQKAREPATLKDSL 62
 DB 359 SAAALAEKGLILDEQAEQGMQAMQOMDAFNQSAEPKQAEVQSR-----QHLEQ 410
 QY 63 DLNNKMFLEKRLPLSGSEAPRLPODPVGM-----RQLOEE-----DEEV 103
 DB 411 SLERIQDRERLQERQGLAD-PEDAAILELNQVALIAELALEILOLOGQGAERLEQL 469


```

QY 104 KARLOPYMAEAEHVLGNNLEGLRQOLKPYTMDLMPVA-----LRVOELQBLR- 152
DB 470 ROELQOLAEQHOAG-ELQRLNGRIA--SLEALQOALDPCGALFMLEQGLQBLRPL 526
QY 153 -----VVCEDTKAQLLGVD----- 167
DB 527 AEGLRVEPCWELAVETVLADLQAVLLDGFGLAGFCGKRLRLSPANGAATAAGSL 586
QY 168 -----EAW-----ALLQGLQSRVYHHTGKRELPHYAESEV-----GIGRHV 206
DB 587 DKVRADADLSPLMLARVKPVETLEQALQARGALDDG-----BSLSRDGYWGRHF 636
QY 207 QOLRSVAPAPASPARLSRCVOVLSRKLTLLKAKALHARI-----QONLDOL 253
DB 637 LVRRRS-----DEAGGGLAAQGL--EALQERRALETRVAEEGERLAARDQRELEGA 690
QY 254 REBELSRAFACTGTEPCGAPDPQMLSEHVQRIOAFRODTYLOIAAFTRAIDQTEVEYOQ 313
DB 691 REQVRROVQEGRRHG-----ELKAQLSA-QQAKVEQOLVLRKRRRLDEEVAELAEQ 739
QY 314 LAPPPEHSFAPEFOOTDSGKVL-----SKLARLDLMDLMDI-TLSLHDQGH 360
DB 740 KALEQEOISEARLTLDLADLSDMLDTERKRSLLAERDALKERLDRIKQDARTH----KDH 755
QY 361 SH 362
DB 796 All 797

RESULT 14
Q98TGI PRELIMINARY: PRT: 259 AA.
AC Q98TGI:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE 28kda-1e apolipoprotein.
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
OX NCBI_TaxID=7937;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA MEDLINE=21175592; PubMed=11278178;
RA Kondo H., Kawazoe I., Nakaya M., Kikuchi K., Aida K., Watabe S.;
RT "The novel sequences of major plasma apolipoproteins in the eel"
RL Biochim Biophys Acta 1531:132-142(2001).
DR EMBL: AB046208; BAB40965.1;
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 1.
KW lipoprotein.
SQ SHQUNCE 259 AA; 29895 MW; C392ADP5BAC71PAF CRC64;

Query Match 8.7%; Score 163; DB 13; Length 259;
Best Local Similarity 22.9%; Pred. No. 0.0042;
Matches 49; Conservative 48; Mismatches 91; Indels 26; Gaps 5;

```

```

QY 176 LOSRVVHHTGFKELFPHYAESLVSGIRRHVEL 209
DB 211 LQKEIFNKIMFDRITLYPYADGLKVKVLDPIYQGL 244

```

```

RESULT 15
Q9GLM8 PRELIMINARY: PRT: 317 AA.
AC Q9GLM8:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Apolipoprotein E.
OS APOE.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.
OX NCBI_TaxID=9593;
RN 111
RP SEQUENCE FROM N.A.
RA Rogaeve E.I., Dvorianchikov G.A., Riazanskaia N.N.;
RT "APOE gene evolution in Homioidae."
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF200502; AAC28579.1;
DR EMBL: AF200500; AAC28579.1; JOINED.
DR EMBL: AF200501; AAC28579.1; JOINED.
DR HSSP: P02649; IOEF.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 1.
KW lipoprotein.
SQ SEQUENCE 317 AA; 36065 MW; 228BPPAH64B52C CRC64;

```

```

Query Match 8.7%; Score 163; DB 6; Length 317;
Best Local Similarity 21.4%; Pred. No. 0.0055;
Matches 67; Conservative 58; Mismatches 120; Indels 68; Gaps 8;

QY 10 WALALLSAFSATQARKGFNDY--FSQTSCKGKGVKQIHQOKMAEPAATLKDSLQDINNMM 67
DB 44 WEAL-----GFFWYLRWVQTLSEVOY-ELLSSQVTOELTALMDETMKELKAY 92
QY 68 NKFL-EKLRPLSGSEAPRLPQDPVGMRLQOELEEVKARLOPYMAEAEHVLGNNLEGLR 126
DB 93 KSRILRQULTVAEETARLSKEIQAAQARIGADMEYKGLAQRGVQAMICOSTEILR 152
QY 127 QOLKPYTMDIMEQVALRVQELQOLRVGSDTKAQLLGVDQAMALLQGLQSRVYHHTGR 186
DB 153 ARLASHIRKLRKRLRLRDADDLQKRLAVYQAGAREGAEV-----SAIRERIGP 201
QY 187 FKHLFPHYASLSVSGIRRHVELHNSVAPAHASPCVQVLSRKLTLLKAKALHARI 246
DB 202 LVEDGRVRAATVGSLSKQPLQER-----AAQKGBRLKARMEGSKRT 243
QY 247 QONLDQLEELSRFAFGTGEAGPDPQMLSEEVROLOAFRODTYLOIAAFTRAI--- 303
DB 244 RDRILDEVKEQVA-----FVRAKILEQAOQIRLOAEARQARHLKSW 282
QY 304 -DOETEEVQOLA 315
DB 283 FEPLVEDMQROMA 295

Search completed: January 23, 2003, 17:45:55
Job time : 79 secs

```